

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO TECNOLÓGICO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

CLEBSON JOEL MENDES DE OLIVEIRA

**MODELO PARA CARACTERIZAÇÃO VISUAL DE CENAS APLICADO À
LOCALIZAÇÃO ROBÓTICA EM AMBIENTES
EXTERNOS DINÂMICOS E AO RECONHECIMENTO DE
OBJETOS**

VITÓRIA
2013

Clebson Joel Mendes de Oliveira

**Modelo para Caracterização Visual de Cenas Aplicado à
Localização Robótica em Ambientes Externos Dinâmicos
e ao Reconhecimento de Objetos**

Dissertação de Mestrado submetida ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Espírito Santo, como requisito parcial para a obtenção do Grau de Mestre em Engenharia Elétrica.

Centro Tecnológico

Programa de Pós-Graduação em Engenharia Elétrica

Universidade Federal do Espírito Santo

Orientador: Profa. Dra. Raquel Frizera Vassallo

Coorientador: Prof. Dr. Evandro Ottoni Teatini Salles

Vitória-ES

2013

Dados Internacionais de Catalogação-na-publicação (CIP)
(Biblioteca Setorial Tecnológica,
Universidade Federal do Espírito Santo, ES, Brasil)

O48m Oliveira, Clebson Joel Mendes de, 1986-
Modelo para caracterização visual de cenas aplicado à
localização robótica em ambientes externos dinâmicos e ao
reconhecimento de objetos / Clebson Joel Mendes de Oliveira. –
2013.

135 f. : il.

Orientador: Raquel Frizera Vassallo.

Coorientador: Evandro Ottoni Teatini Salles.

Dissertação (Mestrado em Engenharia Elétrica) –
Universidade Federal do Espírito Santo, Centro Tecnológico.

1. Robótica. 2. Reconhecimento de padrões. 3.
Probabilidades – Modelos. 4. Mapeamento digital. 5. Visão por
computador. 6. Visão de robô. I. Vassallo, Raquel Frizera. II.
Salles, Evandro Ottoni Teatini. III. Universidade Federal do
Espírito Santo. Centro Tecnológico. IV. Título.

CDU: 621.3

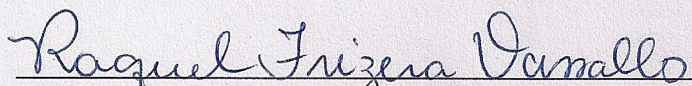
Clebson Joel Mendes de Oliveira

**Modelo para Caracterização Visual de Cenas
Aplicado à Localização Robótica em Ambientes
Externos Dinâmicos e ao Reconhecimento de
Objetos**

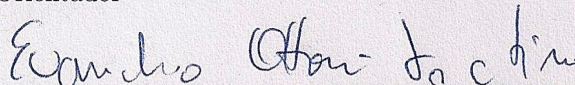
Dissertação submetida ao programa de Pós-Graduação em Engenharia Elétrica do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para a obtenção do Grau de Mestre em Engenharia Elétrica.

Aprovada em 15 de julho de 2013.

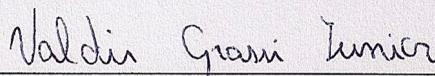
COMISSÃO EXAMINADORA



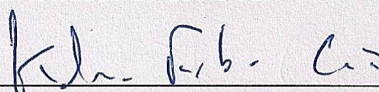
Profa. Dra. Raquel Frizera Vassallo
Universidade Federal do Espírito Santo
Orientador



Prof. Dr. Evandro Ottoni Teatini Salles
Universidade Federal do Espírito Santo
Coorientador



Prof. Dr. Valdir Grassi Jr.
Universidade de São Paulo



Prof. Dr. Klaus Fabian Côco
Universidade Federal do Espírito Santo

Este trabalho é dedicado à minha família:

Ananete, Osvaldo, Amanda, Luana, Marcelo, Luã, Lorenzo e Geise.

Por todo o amor que tem por mim e por todo sacrifício que fizeram para me proporcionar educação de qualidade, formação superior e mestrado.

Agradecimentos

A Deus por ter me dado força e clareza nos momentos em que não era possível contar com o apoio de outras pessoas.

À minha mãe, Ananete Mendes de Oliveira, por ter estado sempre presente ao meu lado, acompanhando todos os passos do trabalho, dividindo angústias e felicidades e por sempre ter me dado amor, independente dos problemas pelos quais estávamos passando.

À minha família, pelo apoio, amor, carinho, compreensão, e por todas as vezes que não estive presente em encontros familiares por conta deste trabalho. Em especial à minha irmã, Luana Mendes de Oliveira, por ter me ajudado a capturar as últimas imagens de teste deste trabalho.

À minha noiva, Geise de Aguiar Quirino, pelo amor, companheirismo e pela enorme compreensão. Por todas as vezes que me forçou a não esquecer da dissertação, mesmo com muito trabalho a fazer. Por todas as vezes que ficamos em casa no fim de semana para eu poder trabalhar na dissertação. Por todas as vezes que dormiu ao som e à luz do meu notebook. Por todas as vezes que deixou de fazer sua dissertação para me ajudar com a minha. E principalmente pelas mais de 3000 imagens de teste que me ajudou a capturar, pois foi ela que segurou a câmera no calor, no frio, no vento e às vezes até no início de chuva. Obrigado meu amor, por tudo.

À minha orientadora, Raquel Frizera Vassallo. Pela orientadora maravilhosa que é, pois me deu liberdade para direcionar o meu trabalho e não me deixou sem assistência. Sempre esteve disponível para conversar, orientar e trabalhar junto no desenvolvimento do projeto. Pela imensa compreensão que teve ao me deixar trabalhar com qualidade, mesmo que isso implicasse na demora para apresentar a dissertação. E por ser amiga antes de ser orientadora, pois em muitos momentos que eu estava imerso no trabalho e só queria conversar sobre a dissertação ela sabiamente mudava o assunto e falava como amiga, me tranquilizando e fazendo com que o trabalho fosse executado de forma mais calma e com mais qualidade.

Ao meu coorientador, Evandro Ottoni Teatini Salles, pela parceria. Por todas as vezes que o atrasei por pará-lo no corredor para tirar dúvidas sobre o trabalho. E acima de tudo pela precisão nas suas orientações, pois sempre soube ouvir e quando falava ia direto ao ponto e rapidamente solucionava minhas dúvidas.

Aos colegas de laboratório do LAI, sempre dispostos a ajudar e dando uma palavra de conforto. Em especial a Vitor Buback Covre, por ter me ajudado a capturar as primeiras imagens de teste deste trabalho.

*Não se preocupe muito com as suas dificuldades em Matemática,
posso assegurar-lhe que as minhas são ainda maiores.
(Albert Einstein)*

Resumo

Considerando os atuais desafios da robótica que estão relacionados à área da cognição, percebe-se que algumas dessas atividades que são simples para seres humanos ainda não são tarefas triviais para robôs. Logo, ainda há vários problemas não resolvidos de forma robusta, que estão relacionados à falha de robôs quando defrontados com o caos do mundo real, principalmente em ambientes externos dinâmicos, tais como: mapeamento, planejamento de trajetória, localização, navegação e identificação de objetos segundo suas formas e funções.

Como muitos desses problemas apresentam chances de serem resolvidos por Visão Computacional, o objetivo deste trabalho foi desenvolver um método de localização visual autônomo *online* aplicado a ambientes externos dinâmicos, sem considerar informações *a priori* através da caracterização visual de cenas.

Para isso, desenvolveu-se um modelo para caracterização visual de cenas baseado na estimação de Funções Massa de Probabilidade (*Probability Mass Functions* - PMFs) de características visuais *Speeded Up Robust Features* (SURF) para os lugares do mapa de um ambiente externo. Utilizando esse modelo para localizar amostras de lugares do mapa, definiu-se um método de localização, que calcula as probabilidades de uma amostra pertencer a lugares candidatos do mapa, e as compara a valores de referência, definidos por curvas *Receiver Operating Characteristic* (ROC) dos lugares candidatos.

Os testes realizados para avaliar o poder de caracterização do modelo desenvolvido, e a qualidade do método de localização proposto utilizaram um mapa visual gerado por um grupo de imagens de caracterização de 28 lugares da Universidade Federal do Espírito Santo (UFES). Já para avaliar a generalização do modelo de caracterização aplicado ao problema de reconhecimento de objetos, foram utilizados grupos de imagens de caracterização de 4 objetos similares.

Pelos resultados obtidos, comprova-se que este trabalho alcançou seus objetivos, pois amostras visuais do ambiente externo dinâmico utilizado foram localizadas corretamente, alcançando desempenho global de classificação pelo menos aceitável ($> 0,7$) segundo o índice *Area Under Curve* (AUC). Para o melhor caso do algoritmo foram gastos 3,361 segundos, e também foi possível reconhecer objetos com desempenho global de classificação pelo menos aceitável, gastando 265 milissegundos no melhor caso.

Palavras-chaves: caracterização visual. modelo probabilístico. ambientes externos.

Abstract

Considering the current Robotics challenges which are related to cognitive problems, one can note that some of these activities easy for human beings to execute are not so easy for robots. Then, there are still some unsolved problems in a robust way, that are related to robots failure caused by the caothic real world, mainly in dynamic outdoor environments such as: mapping , trajectory planning, localization, navigation and object recognition by shape and function.

As the majority of those problems can be solved by Computer Vision techniques, the goal of this work was develop an autonomous online visual localization method applied to dynamic outdoor environments, without considering *a priori* information using visual scene characterization.

For this, it was developed a model for visual scene characterization, based on Probability Mass Functions (PMFs) of visual features SURF from places of an environment map. Using that model to localize place samples of the map, it was defined a localization method, that calculates probabilities of a sample to belong to a candidate place from the map, and compares them to reference values, defined by ROC curves from candidate places.

The tests executed for evaluating the characterization power of the developed model and the quality of the proposed localization method, used a visual map generated from a image group for each 28 places of the environment UFES. For evaluating the generalization of the characterization model aplied to the objects recognition problem, it was used a group of characterization images from 4 similar objects.

Thus, the obtained results show this work reached its goals, because visual samples from the used dynamic outdoor environment were correctly localized , reaching global classification performance at least acceptable ($AUC > 0,7$), spending 3,361 seconds in the best case; and it was also possible to recognize objects from the used group with global classification performance at least acceptable too, spending 265 miliseconds in the best case.

Key-words: visual characterization. probabilistic model. outdoor environments.

Lista de ilustrações

Figura 1 – Modelo gráfico bayesiano para localização de robôs móveis. Fonte: Thrun, Burgard e Fox (2005).	21
Figura 2 – Grau de desenvolvimento das tecnologias de localização em relação à precisão alcançada pelas mesmas. Fonte: Hazas, Scott e Krumm (2004).	22
Figura 3 – Robô explorador de Marte que utiliza doze câmeras para se localizar. Fonte: NASA (2012).	24
Figura 4 – Cálculo da soma de intensidade de pixels em uma região retangular usando imagens integrais. Fonte: Bay <i>et al.</i> (2008).	33
Figura 5 – Da esquerda para direita tem-se a aproximação utilizada no método SURF respectivamente para derivada parcial de segunda ordem da Gaussiana na direção y (D_{yy}) e xy (D_{xy}). As regiões em cinza apresentam valores iguais a zero. Fonte: Bay <i>et al.</i> (2008).	34
Figura 6 – A cabana dentro do círculo laranja representa o objeto de interesse. À esquerda temos a cabana em uma escala menor dentro da cena e à direita, após uma operação de <i>zoom in</i> , temos a cabana em uma escala maior.	35
Figura 7 – Aplicação do espaço de escalas. Ao invés de reduzir o tamanho da imagem (esquerda) as imagens integrais são utilizadas para aumentar a escala do filtro (direita) a um custo computacional constante. Fonte: Bay <i>et al.</i> (2008).	35
Figura 8 – Pontos de interesse detectados em um campo de girassóis pelo método detector <i>Fast-Hessian detector</i> . Fonte: Bay <i>et al.</i> (2008).	36
Figura 9 – Filtros <i>Haar Wavelet</i> para calcular as respostas na direção x (esquerda) e na direção y (direita). As partes pretas do filtro apresentam peso -1 e as partes brancas $+1$. Fonte: Bay <i>et al.</i> (2008).	38
Figura 10 – Determinação da orientação de um ponto de interesse: uma janela deslizante de orientação de tamanho $\pi/3$ determina a orientação dominante das respostas <i>Haar Wavelet</i> suavizada por uma Gaussiana, definidas para cada ponto de amostragem dentro de uma vizinhança circular em torno do ponto de interesse. Fonte: Bay <i>et al.</i> (2008).	38
Figura 11 – Imagem que mostra o tamanho da janela das regiões orientadas em diferentes escalas. Fonte: Bay <i>et al.</i> (2008).	39

Figura 12 – Para construir o descritor, é colocado sobre o ponto de interesse um quadrado orientado dividido em 4x4 sub-regiões quadradas (esquerda). Nas divisões 2x2 (direita) das sub-regiões 4x4 são calculadas as respostas <i>Haar Wavelet</i> , d_x e d_y , relativamente à orientação do ponto. Para montar o descritor v , são computadas as somas $\sum dx$, $\sum dx $, $\sum dy$, e $\sum dy $ para cada sub-região 4x4. Fonte: Bay <i>et al.</i> (2008).	40
Figura 13 – Se o contraste entre dois pontos de interesse é diferente, esses candidatos a correspondentes são descartados. Fonte: Bay <i>et al.</i> (2008).	41
Figura 14 – Avaliação de desempenho das versões do detector SURF, FH-9 e FH-15, através do critério de avaliação de repetibilidade considerando respectivamente as seguintes características das imagens: mudança de ponto de vista (acima à esquerda), mudança de escala (acima à direita) e adição de níveis de borrão (abaixo). Fonte: Bay <i>et al.</i> (2008).	42
Figura 15 – Curvas <i>recall/precision</i> geradas pelo algoritmo de <i>matching</i> baseado no vizinho mais próximo (<i>Nearest Neighbor</i>), considerando variações dos descritores SURF tanto para as versões padrão (à esquerda), quanto para suas versões estendidas (à direita). Fonte: Bay <i>et al.</i> (2008).	43
Figura 16 – Comparação entre curvas ROC de classificadores <i>Naive Bayes</i> utilizando o modelo para representação de objetos <i>bag-of-words</i> , aplicado a diferentes métodos descritores. Os pontos de interesse utilizados pelos métodos descritores foram gerados de duas formas: através da determinação aleatória de píxeis de borda (à esquerda), e pelo detector SURF (à direita). Fonte: Bay <i>et al.</i> (2008).	44
Figura 17 – À esquerda correspondências corretas e à direita correspondências erradas.	46
Figura 18 – Exemplo de ambiente externo e seus lugares L_i representados pelos marcadores amarelos. Observe que L_i é a i -ésima realização da variável aleatória L	48
Figura 19 – Exemplo de amostra completa para uma caneca.	51
Figura 20 – Exemplo de cena a ser caracterizada.	52
Figura 21 – Comparação entre imagens da mesma cena, porém capturadas por câmeras com FOVs diferentes. À esquerda pode-se observar a imagem obtida por uma câmera com aproximadamente 180° de FOV, e à direita é apresetada uma imagem que simula a diminuição do FOV da câmera.	53
Figura 22 – Exemplo de função de frequência absoluta das características de um lugar.	54
Figura 23 – Exemplo de determinação de limiar utilizado para testar a localização de uma amostra em lugar qualquer do mapa. O ponto circulado sobre a curva representa o limiar que gera melhores taxas de TPR e FPR.	57
Figura 24 – Exemplo de amostra completa para uma controle remoto.	58

Figura 25 – Representação gráfica do espaço ROC. Quanto mais próximo do canto superior esquerdo do gráfico, melhor é o desempenho do classificador associado ao ponto. Fonte: Fawcett (2006).	61
Figura 26 – Na Figura (a) é possível comparar as curvas sem utilizar a área sob elas. No entanto, em (b), a métrica AUC mostra-se necessária para que seja possível comparar as duas curvas. Fonte: Fawcett (2006).	62
Figura 27 – Imagens de exemplo dos lugares do mapa.	67
Figura 28 – Curvas ROC dos lugares e do mapa, geradas por 3 amostras completas de caracterização e 13 amostras simples (1 imagem apenas) de teste por lugar, totalizando 65 testes por lugar e 325 para o mapa.	69
Figura 29 – Mapa do ambiente UFES, do qual foram utilizados, neste experimento, 27 dos seus 28 lugares representados pelos marcadores amarelos.	72
Figura 30 – Imagens exemplo das amostras de caracterização dos lugares do mapa.	76
Figura 31 – Imagens exemplo das amostras de teste dos lugares do mapa.	78
Figura 32 – Curvas ROC dos lugares e do mapa, geradas por 10 amostras completas de caracterização e 12 amostras completas de teste por lugar, totalizando 324 testes por lugar e 8748 para o mapa.	84
Figura 33 – À esquerda temos imagens exemplo das amostras de caracterização e à direita de teste, correspondentes respectivamente de cima para baixo aos lugares L7, L8 e L11.	88
Figura 34 – Curvas ROC dos lugares e do mapa, geradas por 10 amostras completas de teste por lugar, totalizando 270 testes por lugar e 7290 para o mapa.	96
Figura 35 – Imagens exemplo das amostras de teste incompletas dos lugares do mapa.	102
Figura 36 – Curvas ROC dos lugares e do mapa, geradas por 16 amostras incompletas (2 imagens apenas) de teste por lugar, totalizando 448 testes por lugar e 12544 para o mapa.	109
Figura 37 – Comparação das probabilidades das amostras em torno do lugar L23, para diferentes limiares utilizados para descartar características da PMF de L23. Em azul a curva l0 de limiar 1, em verde a curva l1 de limiar 8, e em vermelho a curva l2 de limiar 10.	111
Figura 38 – Comparação das curvas ROC do mapa com amostras teste completas, para três limiares de descarte das características das PMFs. Em azul a curva l0 de limiar 1, em verde a curva l1 de limiar 8, e em vermelho a curva l2 de limiar 10.	113
Figura 39 – Comparação das curvas ROC do mapa com amostras de teste incompletas, para três limiares de descarte das características das PMFs. Em azul a curva l0 de limiar 1, em verde a curva l1 de limiar 8, e em vermelho a curva l2 de limiar 10.	114

Figura 40 – Probabilidades das amostras em relação a L23, acima temos os 5 experimentos e abaixo a média destes experimentos.	116
Figura 41 – Probabilidades das amostras em relação a L22, acima temos os 5 experimentos e abaixo a média destes experimentos.	117
Figura 42 – Média das probabilidades dos 5 experimentos, em relação a L23 e a L22.	118
Figura 43 – Imagens exemplo das amostras de caracterização dos objetos.	119
Figura 44 – Imagens exemplo das amostras de teste dos objetos.	119
Figura 45 – Curvas ROC dos objetos e do classificador, geradas por 5 amostras completas (2 imagens apenas) de teste por objeto, totalizando 20 testes por objeto e 80 para o classificador.	120
Figura 46 – Comparação entre as curvas ROC do classificador. Em azul a curva l0 de limiar 1, em verde a curva l1 de limiar 2, e em vermelho a curva l2 de limiar 3.	121
Figura 47 – Código de teste das bibliotecas de medição de tempo.	123
Figura 48 – Gráfico de utilização dos recursos do sistema durante a execução dos algoritmos de mapeamento de localização. Para todos os teste executados, os gráficos gerados pelo monitor do sistema apresentaram as mesma características, tais quais as apresentadas nesta figura.	125
Figura 49 – Câmera.	134
Figura 50 – Notebook 1.	134
Figura 51 – Notebook 2.	135
Figura 52 – Notebook 3.	135

Lista de tabelas

Tabela 1	– Exemplo de ocorrência de características em 3 amostras de um lugar. .	54
Tabela 2	– Correspondência das características da amostra c capturada, com as características s_i do lugar L_i descrito pela Tabela 1.	56
Tabela 3	– Exemplo de matriz de confusão.	64
Tabela 4	– Relação entre o índice AUC e o número de amostras completas de caracterização.	68
Tabela 5	– Quantidade de características extraídas dos lugares do mapa.	70
Tabela 6	– Matriz de confusão* do mapa.	71
Tabela 7	– Períodos de captura das amostras de teste e de caracterização, utilizadas no experimento. Na segunda e terceira colunas temos o dia, a hora e os minutos de início e fim da captura, e na quarta coluna temos o intervalo de tempo mínimo (em meses) entre a captura das imagens de caracterização e das imagens de teste.	73
Tabela 8	– Quantidade de características extraídas dos lugares do mapa.	85
Tabela 9	– Matriz de confusão* dos lugares para amostras de teste completas. . .	87
Tabela 10	– Quantidade de características extraídas dos lugares do mapa antes e após a inversão das imagens de teste com as imagens de caracterização. .	97
Tabela 11	– Matriz de confusão* dos lugares para amostras de teste completas após a inversão de imagens de teste com imagens de caracterização de alguns lugares do mapa.	98
Tabela 12	– Períodos de captura das amostras de teste e de caracterização, utilizadas no experimento com amostras incompletas. Na segunda e terceira colunas temos o dia, a hora e os minutos de início e fim da captura, e na quarta coluna temos o intervalo de tempo mínimo (em meses) entre a captura das imagens de caracterização e das imagens de teste.	103
Tabela 13	– Número de características da PMF de L23, e tempo de médio de cálculo das probabilidades das amostras de teste pertencerem ao lugar L23 para os limiares 1, 8 e 10, do algoritmo de descarte de características. .	111
Tabela 14	– Tempo de médio de localização das amostras de teste no melhor caso, tanto para o experimento com amostras completas quanto incompletas, para os limiares 1, 8 e 10, do algoritmo de descarte de características da PMF dos lugares do mapa.	112
Tabela 15	– Matriz de confusão* dos objetos, para 5 amostras de teste completas. .	121
Tabela 16	– Tempo de médio de classificação das amostras de teste no melhor caso, em milissegundos, para os limiares 1, 2 e 3, do algoritmo de descarte de características da PMF dos objetos.	121

Tabela 17 – Tempo de médio de localização/classificação das amostras de teste no melhor caso, e de construção do mapa para os principais experimentos apresentados no Capítulo 5.	125
---	-----

Lista de abreviaturas e siglas

ROV	<i>Remotely Operated Vehicle</i>
PMF	<i>Probability Mass Function</i>
SIFT	<i>Scale-invariant Feature Transform</i>
SURF	<i>Speeded Up Robust Features</i>
RANSAC	<i>Random Sample Consensus</i>
SLAM	<i>Simultaneous Localisation and Mapping</i>
FOV	<i>Field Of View</i>
ROC	<i>Receiver Operating Characteristic</i>
AUC	<i>Area Under Curve</i>
UFES	Universidade Federal do Espírito Santo

Lista de símbolos

σ	Desvio padrão de uma função gaussiana
e	Valor da escala de um ponto de interesse no método SURF
L	Variável aleatória utilizada para representar o processo de localização de um grupo de imagens quaisquer em um mapa visual
E	Experimento estatístico que representa a determinação da localização de um grupo de imagens quaisquer
ξ	Possível resultado do experimento estatístico E
f	Função identidade
Pr	Probabilidade
PMF_L	Função massa de probabilidade da variável aleatória L
L_i	i -ésima realização da variável aleatória L , ou seja, um lugar do mapa visual
s_i	Conjunto de características extraídas de um grupo de imagens utilizado para caracterizar o lugar L_i
PMF_{s_i}	Função massa de probabilidade da variável aleatória s_i , que representa o processo de localização de um conjunto de características em um mapa visual
FA	Função estatística de frequência absoluta
FR	Função estatística de frequência relativa
s_{iw}	Característica qualquer do conjunto s_i
k	Quantidade de características presentes no conjunto s_i
c	Conjunto de características extraídas de um grupo de imagens capturadas em um lugar qualquer
c_j	Característica qualquer do conjunto c
m	Quantidade de características presentes no conjunto c

Sumário

I	INTRODUÇÃO	19
1	INTRODUÇÃO	20
1.1	O Problema	21
1.2	Revisão Bibliográfica	25
1.3	Objetivos	27
1.4	Contribuições	28
1.5	Estrutura da Dissertação	28
II	FUNDAMENTAÇÃO TEÓRICA	29
2	EXTRAÇÃO DE CARACTERÍSTICAS	30
2.1	Imagens Naturais	30
2.2	Método SURF	31
2.2.1	Detector de Ponto de Interesse	32
2.2.2	Descritor de Ponto de Interesse e Determinação de Correspondências	36
2.2.2.1	Passo 1: determinação da orientação	37
2.2.2.2	Passo 2: extração de informação	39
2.2.2.3	Passo 3: determinação de correspondências	40
2.2.3	Avaliação do método SURF	41
III	MODELAGEM	45
3	CARACTERIZAÇÃO DE CENAS EM AMBIENTES EXTERNOS	46
3.1	Caracterização de Cenas Aplicada ao Problema de Localização: O Modelo Proposto	47
3.1.1	Caracterização de Lugares	47
3.1.2	Determinação de Correspondências entre Lugares	55
3.2	Caracterização de Cenas Aplicada ao Problema de Reconhecimento de Objetos	58
IV	EXPERIMENTOS	59
4	METODOLOGIA DE AVALIAÇÃO	60
4.1	Curvas ROC	60
4.2	Matriz de Confusão	64

5	AVALIAÇÃO DO MODELO	66
5.1	Número de Imagens para Caracterização	66
5.2	O Problema de Localização	71
5.2.1	Análise Geral	78
5.2.2	Influência da Quantidade de Características	84
5.2.3	Análise dos Lugares com Desempenho Abaixo de Aceitável	88
5.2.3.1	L7, L8 e L11	88
5.2.3.2	L0, L3 e L10	89
5.2.3.3	L12 e L13	89
5.2.3.4	L6 e L15	90
5.3	Influência da Iluminação	90
5.4	Utilização de Amostras de Teste Incompletas	99
5.5	Descarte de Características da PMF dos Lugares	109
5.6	Comportamento em Regiões de Transição entre Lugares	114
5.7	O Problema de Reconhecimento de Objetos	118
6	CARACTERÍSTICAS TEMPORAIS	122
6.1	Métricas para Medição de Tempo	122
6.2	Análise de Tempo de Resposta	123
6.3	Tempos de Resposta dos Experimentos	125
7	CONCLUSÃO	127
	Referências	129
	ANEXOS	133
	ANEXO A – EQUIPAMENTOS UTILIZADOS	134

Parte I

Introdução

1 Introdução

Os avanços alcançados pela computação e pela eletrônica na última década, promoveram o desenvolvimento de tecnologias que afetam cada vez mais o nosso cotidiano. Os principais exemplos disso são a Computação em Núvens ([ARMBRUST et al., 2009](#); [MELL; GRANCE, 2011](#); [APPLE, 2012](#)), que possibilita o acesso a informações pessoais através da internet e com alta disponibilidade; e a Robótica de Serviço, tanto industrial ([ABB, 2012](#)) quanto doméstica ([ROBOTSHOP, 2012](#)), que auxilia desde pessoas em seus serviços domésticos até funcionários em uma indústria. Essas tecnologias facilitam cada vez mais a vida dos seres humanos, tanto pela cooperação com estes, por exemplo os *Remotely Operated Vehicles* (ROVs) ([OCEANEERING, 2012](#)) utilizados em campos de petróleo de águas profundas, quanto substituindo-os em serviços repetitivos e maçantes como atividades domésticas ([ADOLF, 2012](#)).

Porém, percebe-se que algumas atividades simples para seres humanos ainda não são tarefas triviais para robôs, ([KYRKI; KRAGIC, 2011](#)). Ainda há vários problemas não resolvidos de forma robusta, que estão relacionados à falha de robôs quando defrontados com o caos do mundo real ([SPERO, 2004](#); [KYRKI; KRAGIC, 2011](#)), principalmente em ambientes externos¹ dinâmicos, tais como: mapeamento, planejamento de trajetória, localização, navegação e identificação de objetos segundo suas formas e funções.

Muitos desses problemas apresentam chances de serem resolvidos por Visão Computacional, que é parte integrante do estado da arte de robótica ([KYRKI; KRAGIC, 2011](#)), uma vez que a percepção visual é uma tecnologia muito flexível, já que a mudança da aplicação implica apenas na mudança das técnicas utilizadas sobre a imagem e não necessariamente altera a câmera. Por isso, Visão Computacional pode ser utilizada para resolver problemas de diversas áreas, desde industrial até robótica assistiva.

O avanço das técnicas de visão tem ajudado a resolver problemas desde manipulação e teleoperação até localização e mapeamento ([KYRKI; KRAGIC, 2011](#)). Dentre esses problemas, um dos mais explorados por pesquisadores é a navegação de robôs móveis, que consiste em fazer com que um robô possa se mover em um ambiente, objetivando a realização de tarefas, como por exemplo: vigilância, exploração e resgate, assistência a pessoas, serviços domésticos, serviços agrícolas, entre outras. Para resolver o problema de navegação é necessário resolver alguns subproblemas, que segundo [Spero \(2004\)](#) podem ser vistos como: mecanismo de locomoção, sistema de controle, sensoriameto, mapeamento do ambiente, planejamento de trajetória e localização.

¹ Neste texto as palavras externo(s) e externa(s) serão utilizadas como tradução livre da palavra *outdoor* do inglês.

1.1 O Problema

A localização de um robô móvel é uma instância do problema geral de localização, que é um problema de percepção fundamental em Robótica. Praticamente todas as tarefas realizadas em Robótica requerem o conhecimento da localização do objeto de interesse, seja este um objeto a ser transportado pelo robô ou o próprio robô.

A Figura 1 mostra um modelo para o problema de localização de robôs móveis. O objetivo da localização é inferir a postura do robô a cada instante de tempo t , ou seja, x_t dado um mapa (m), a percepção do ambiente (z) e as ações de controle (u) (THRUN; BURGARD; FOX, 2005). Nesse modelo considera-se respectivamente x_{t-1} , x_t , x_{t+1} como a postura do robô no instante de tempo anterior, no instante de tempo atual e no instante de tempo futuro.

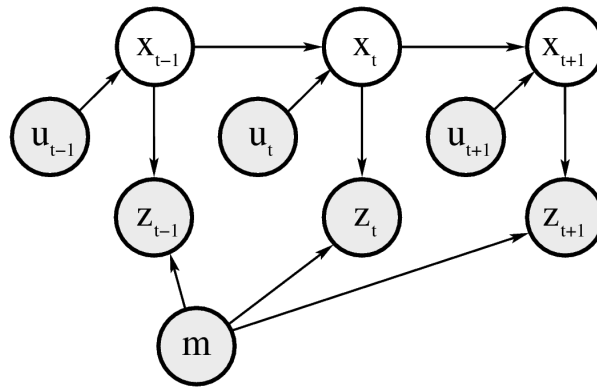


Figura 1 – Modelo gráfico bayesiano para localização de robôs móveis. Fonte: Thrun, Burgard e Fox (2005).

A localização de robôs móveis segundo a abordagem métrica, pode ser vista como um problema de transformação de coordenadas. Os mapas são descritos em um sistema de coordenadas globais, que é independente da postura do robô, e a localização é o processo que estabelece correspondências entre o sistema de coordenadas do mapa e o sistema de coordenadas local do robô. Então, é possível determinar a localização do robô expressando sua postura no mesmo sistema de coordenadas do mapa (THRUN; BURGARD; FOX, 2005).

Todavia, nem todos os problemas de localização podem ser resolvidos pela abordagem métrica, pois pode não se dispor de um sistema global de localização. Por exemplo, quando um robô precisa executar tarefas em ambientes externos dinâmico, como ambientes urbanos, nem sempre é possível obter informações métricas confiáveis do ambiente, pois este pode mudar bastante durante a execução da tarefa.

Nessa situação podem ser utilizadas as abordagens topológica e por aparência, pois estas utilizam respectivamente a forma e informações implícitas do ambiente e seus

objetos, tais como bordas, fluxo de superfícies e pontos de interesse.

Considerando o contexto da navegação de robôs móveis, pode-se observar que os problemas de localização e mapeamento estão diretamente ligados e são utilizados conjuntamente para resolver o problema da navegação autônoma.

Para tal, várias tecnologias podem ser utilizadas e algumas destas são apresentadas na Figura 2, que está um pouco desatualizada em relação à precisão das tecnologias, se considerarmos sua data de publicação. Assim, de forma geral, todas estas tecnologias podem ser deslocadas para a esquerda com o objetivo de atualizar a Figura 2.

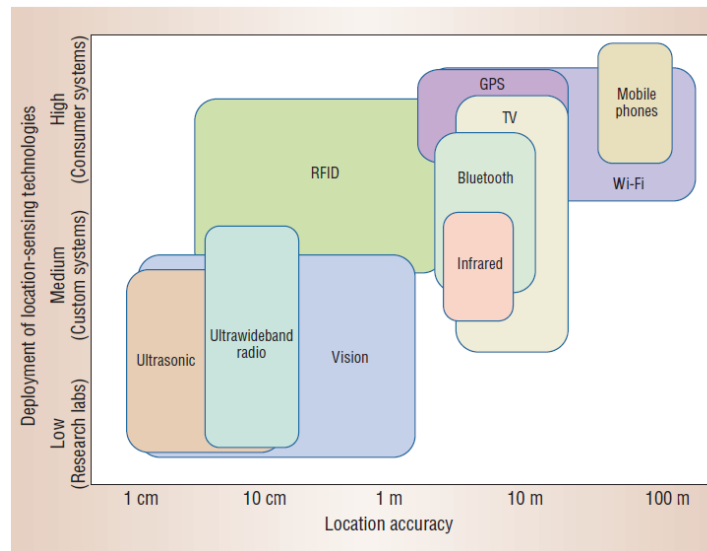


Figura 2 – Grau de desenvolvimento das tecnologias de localização em relação à precisão alcançada pelas mesmas. Fonte: Hazas, Scott e Krumm (2004).

Dentre as tecnologias apresentadas na Figura 2, uma bastante utilizada para localização em ambientes externos é a *Global Position System* (GPS) que se baseia em sinais de satélite. Esta tecnologia tem se tornado cada vez mais barata e acessível, porém ainda existem muitas variáveis de influência que afetam a precisão dos sistemas baseados em GPS tais como: *clock bias*, *clock drift*, ionosfera, reflexão por vários caminhos, disponibilidade dos satélites, algoritmo computacional de tratamento do sinal, entre outras (CHEN, 2013).

Existem várias técnicas para solucionar o problema da precisão dos sistemas baseados em sinais de satélite, tais como *Carrier-Phase Enhancement* (CP) e *Real Time Kinematic* (RTK), que na teoria conseguem reduzir a precisão desses sistemas para ordem de centímetro. Em Jensen *et al.* (2012) foram comparados os sistemas convencionais de GPS com sistemas RTK-GPS.

Neste, os testes de campo utilizando GPS mostraram que mais de 95% dos erros de posição estavam dentro do intervalo de 0 a 2,6 metros. Já os testes de campo utilização

RTK-GPS mostraram que mais de 95% dos erros de posição estavam dentro do intervalo de 0 a 0,2 metros. Outro experimento conduzido por um grupo de pesquisa na Coreia aplicou RTK-GPS em um carro de passeio e conseguiu 20 centímetros de valor *Root Mean Square* (RMS) para o erro de posição (CHEN, 2013).

Entretando, como discutido em Chen (2013), testes baseados em RTK-GPS para localizar veículos industriais, realizados pela Scania AB em 2012 na Suécia mostraram erros maiores, pois os valores RMS dos erros de posição foram de 0,67 metros na direção norte 0,32 metros na direção leste.

Os valores de precisão alcançados por sistemas baseados em GPS, que foram apresentados anteriormente, mostram que é possível obter precisão abaixo de 1 metro, e em algumas aplicações este valor pode ser considerado excelente. No entanto, os equipamentos baseados em GPS que são facilmente encontrados no mercado a preços acessíveis, não apresentam boa precisão abaixo de 10 metros como mostra a Figura 2, por não implementarem os métodos de melhoramento de medição tais como RTK-GPS.

Além disso, equipamentos baseados RTK-GPS são muito caros se comparados aos respectivos equipamentos que não utilizam a tecnologia RTK (TAKASU; YASUDA, 2009), apresentando custo de varejo por equipamento entre U\$ 15.000,00 e U\$ 60.000,00 (GRISSE; ALLEY; HEATWOLE, 2009).

Assim, justifica-se a busca pelo desenvolvimento de outros métodos para efetuar localização em ambientes externos, tais como métodos baseados em Visão Computacional, que estão em crescente utilização atualmente.

A utilização de métodos de Visão aplicados ao problema de localização em ambientes externos não necessariamente precisa ser uma alternativa à tecnologia GPS, uma vez que a cooperação desses métodos pode aumentar a robustez de sistemas de localização, pois na ausência do sinal de GPS o sistema de Visão pode resolver o problema, e em algumas aplicações o sistema de Visão pode aumentar a precisão do sistema baseado em GPS, como por exemplo, a manipulação de objetos em um ambiente.

Além disso, existem certas aplicações que não podem utilizar sistemas baseados em GPS, simplesmente pela ausência desse sistema no ambiente, como por exemplo a exploração extraterrestre. Nesse contexto, a *National Aeronautics and Space Administration* (NASA) tem desenvolvido robôs equipados com câmeras para explorar a superfície de planetas, tais como o *All-Terrain Hex-Limbed Extra-Terrestrial Explorer* (ATHLETE), Figura 3, que não utilizam sistema de GPS para efetuar localização.



Figura 3 – Robô explorador de Marte que utiliza doze câmeras para se localizar. Fonte: [NASA \(2012\)](#).

Logo, a utilização de Visão Computacional para resolver o problema da localização em ambientes externos mostra-se de grande valia. Entretanto, a alta precisão conseguida com técnicas de Visão Computacional exposta na Figura 2 é comumente alcançada em ambientes internos² e a aplicação destas técnicas em ambientes externos não é direta.

Por isso, algumas das ferramentas existentes e amplamente aplicadas em ambientes internos, não funcionam ou não apresentam um bom desempenho em ambientes externos dinâmicos, pois estes apresentam características que tornam o processo de navegação bem mais complexo do que em ambientes internos, tais como: ausência de marcos artificiais pré-definidos, desníveis no terreno, variação de iluminação e variabilidade de objetos no ambiente, como discutido em [Spero \(2004\)](#).

Dessa forma, é necessário utilizar estruturas de ocorrência natural nos ambientes externos para alcançar desempenho similar à navegação autônoma em ambientes internos. Portanto, é importante desenvolver algoritmos confiáveis que sejam capazes de extrair características de um grande conjunto de dados ruidosos para que seja possível o desenvolvimento de aplicações de navegação autônoma em ambientes externos ([ZHANG; XIAO; XIE, 2008](#)).

Uma abordagem atualmente utilizada para tratar esse tipo de problema é a utilização de métodos de Visão Computacional combinados com modelos probabilísticos. As técnicas para sensoramento visual devem ser as menos variantes possíveis às condições do ambiente, pois em ambientes externos essas variam consideravelmente, como por exemplo a variação de iluminação, tornando assim o sistema mais suscetível a falhas. Já o tratamento das características do ambiente por modelos probabilísticos é justificado pelo ganho de robustez alcançado por tais métodos, uma vez que estes conseguem tratar as incertezas

² Neste texto as palavras interno(s) e interna(s) serão utilizadas como tradução livre da palavra *indoor* do inglês.

intrínsecas ao processo de sensoriamento e atuação de robôs (THRUN; BURGARD; FOX, 2005).

Dessa forma, será exposta a seguir uma breve revisão bibliográfica sobre as principais técnicas de localização visual, desde aquelas eficazes somente em ambientes internos até as desenvolvidas para ambientes externos.

1.2 Revisão Bibliográfica

Os primeiros avanços dos artigos descritos a seguir foram estruturados sobre métodos de reconhecimento de lugares através de mapas baseados em *landmarks*, na sua maioria aplicados em ambientes internos. Um método de localização baseado em *landmarks* para ambientes internos foi apresentado em Dudek e Jugessur (2000), no qual as características são estatisticamente determinadas e comparadas usando análise de componente principal no domínio da frequência. Através da utilização dessa técnica os objetos podem ser identificados sobre *backgrounds* e orientações variáveis. Assim, o objeto pode servir de *landmark*. No entanto, essa abordagem tenta reconhecer objetos isolados que devem ser analisados e treinados através de dados preparados e selecionados manualmente.

Em Ulrich e Nourbakhsh (2000) é apresentada uma abordagem para navegação que usa mapas topológicos combinados a histogramas de cor HLS (*Hue, Lightness and Saturation*) e RGB (*Red, Green and Blue*) gerados das imagens capturadas em cada nó do mapa, e então, é aplicado o método de classificação *Nearest Neighbor Learning*. Apesar da abordagem ter apresentado um desempenho bom, os experimentos apresentados no entanto, focaram em ambientes estruturados internos ou em ambientes urbanos, os quais apresentam grandes quantidades de *landmarks* artificiais, como prédios por exemplo.

Em Boada, Blanco e Moreno (2004), é apresentada uma forma de organizar mapas topológicos obtidos através de diagramas de Voronoi e também de reconhecer lugares através das características desses diagramas. Contudo, as considerações básicas utilizadas para caracterizar os lugares do diagrama (nós do mapa) são baseadas em formas de corredor ou ambientes com aparência de sala, ou seja, ambientes internos.

Uma segunda grande contribuição foi a criação de métodos de extração de características do ambiente baseados em imagens, que são invariantes ou pouco variantes às condições do ambiente, pois esse tipo de ferramenta foi o passo inicial para se trabalhar em ambientes externos.

Nestes é inviável a utilização direta de métodos de localização até então difundidos para ambientes internos, como os algoritmos *Extended Kalman Filter Localization* (*EKF Localization*), *Unscented Kalman Filter Localization* (*UKF Localization*), *Grid Localization* e *Monte Carlo Localization* (THRUN; BURGARD; FOX, 2005), pois o principal problema

da utilização de mapas baseados em informações métricas aplicados a ambientes externos está no fato de que ambientes externos são altamente dinâmicos, e assim o local anteriormente mapeado e considerado livre pode estar ocupado ou pode ter sido modificado significativamente.

Assim, considerando métodos de extração de características visuais baseados em pontos de interesse ou características (*features*), os trabalhos de Lowe e Little (2001), e Andreasson e Duckett (2004) utilizaram *Scale-invariant Feature Transform* (SIFT) (LOWE, 1999) para extrair de imagens características do ambiente, que foram utilizadas como *landmarks* no algoritmo de localização do robô. Porém, apesar deste método incorporar robustez ao processo de localização, pela utilização de extratores de características invariantes às condições de aquisição das imagens do ambiente (rotação, translação e iluminação), a aplicação desenvolvida ainda é restrita a ambientes internos durante a maior parte do tempo.

A criação do método SIFT (LOWE, 1999) e sua posterior melhoria (LOWE, 2004) promoveram um grande avanço na área de Visão Computacional/Visão Robótica, pela sua alta robustez à variação das características do ambiente, que foi primeiramente alcançada por este método. Porém, o fato do custo computacional desse método ser muito elevado dificultou sua aplicação em sistemas autônomos *online*, como por exemplo a navegação robótica autônoma em ambientes externos.

Poucos anos após a publicação da versão final do SIFT (LOWE, 2004), foi criado o método *Speeded Up Robust Features* (SURF) (BAY; TUYTELAARS; GOOL, 2006; BAY *et al.*, 2008), que conseguiu superar a relação custo benefício do SIFT por apresentar excelentes resultados em relação à acurácia, às vezes superando o SIFT, e um custo computacional viável que permitiu e ainda permite sua aplicação em sistemas autônomos *online*.

A partir desse ponto, a aplicação do método de extração de características visuais SURF foi amplamente difundida nas áreas de Robótica e Visão Computacional, e ajudou a promover avanços na área de robótica autônoma *online* em ambientes externos, como exemplificam os artigos a seguir que foram publicados entre 2008 e 2011.

Em Valgren e Lilienthal (2008) são construídos mapas topológicos probabilísticos baseados em aparência, ou seja, sensores de visão implementados através da aquisição de imagens panorâmicas e a posterior extração de características pelo método SURF. Para tornar o desempenho satisfatório são também aplicados os modelos de geometria epipolar (*epipolar constraint*) e *Random Sample Consensus* (RANSAC). Porém, os resultados apresentados foram gerados por uma base de dados de imagens não exclusivamente de ambientes externos, e o método não foi aplicado em um sistema robótico autônomo *online*.

Uma das contribuições atuais mais bem sucedidas foi a formulação de um novo

modelo de Mapeamento e Localização Simultâneos (SLAM) baseado apenas em aparência (imagens omni-direcionais e estéreo), o algoritmo *Fast Appearance-Based Mapping* (FAB-MAP) (CUMMINS; NEWMAN, 2008), e sua posterior melhoria FAB-MAP 2.0 (CUMMINS; NEWMAN, 2010), pois foram testados em navegação em grandes distâncias.

A principal contribuição destes trabalhos foi a realização de dois testes em ambientes externos: um de 7 km de rota e outro de 1000 km. Imagens foram capturadas ao longo das distâncias descritas e a base de dados gerada foi utilizada para avaliar o modelo. Apesar de ambos os testes terem sido bem sucedidos, considerando as distâncias nunca efetuadas anteriormente, esse algoritmo é dependente de modelo e precisa ser treinado, pois é baseado em aprendizagem. O treinamento utilizou 1921 imagens que produziram 2,5 milhões de descritores SURF com vizinhança de tamanho 128 que ocuparam 3,2 GB. Além disso, o algoritmo de localização utiliza informações de instantes anteriores para estimar a localização atual; as imagens foram capturadas durante apenas seis dias, caracterizando assim pouca chance de mudança no ambiente; e não foram efetuados testes *online*.

Por fim, em Mitsuhashi, Sakai e Kuroda (2010) e Mitsuhashi e Kuroda (2011) são desenvolvidos sistemas para localização de robôs móveis através de reconhecimento de lugares baseados em aparência através do método SURF. O algoritmo de *matching* é feito através da comparação entre a imagem obtida e uma imagem de referência do local no mapa. Todos os testes realizados foram em ambientes externos e o erro de posição médio obtido foi menor que 2% sem a utilização de GPS. Contudo, os resultados só foram alcançados através da fusão de sensores e da utilização do algoritmo FAB-MAP (CUMMINS; NEWMAN, 2008) já citado anteriormente.

1.3 Objetivos

Dessa forma, considerando a navegação externa de robôs móveis autônomos através de características visuais, o objetivo deste trabalho é desenvolver um método de localização visual autônomo *online* aplicado a ambientes externos dinâmicos, sem considerar informações *a priori*, através da caracterização visual de cenas. Pois, uma vez que o algoritmo baseia-se apenas nos dados extraídos do ambiente, não é dado como entrada nenhum conhecimento ou estimativa da localização da câmera em instantes anteriores, e não são utilizados *landmarks*.

Logo, pretende-se que o método proposto sirva de base para várias aplicações em robótica visual autônoma *online* em ambientes externos dinâmicos, tais como: navegação de robôs terrestres ou aéreos, reconhecimento de lugares e de objetos.

Além disso, como o método de localização proposto não utiliza a localização da câmera em instantes anteriores, este pode ser acelerado se for estendido para utilizar informações de navegação, e assim passar a utilizá-la.

O método de localização proposto é baseado no método de extração e descrição de características visuais SURF (BAY *et al.*, 2008), aplicado a imagens panorâmicas obtidas por uma única câmera perspectiva com lente grande angular. Para tal, são utilizados conceitos de Robótica Probabilística e Visão Computacional, para que os métodos de navegação que venham a utilizar o método de localização proposto neste trabalho, possam ser aplicados de forma robusta em ambiente externos e dinâmicos.

1.4 Contribuições

As contribuições do presente trabalho são as seguintes:

- desenvolvimento de um modelo para caracterização de ambientes dinâmicos, que pode ser aplicado em ambientes externos;
- desenvolvimento de um método de localização robótica autônoma, que pode ser aplicado em ambientes externos dinâmicos de forma *online*;
- desenvolvimento de uma metodologia para avaliar métodos de localização em ambientes externos;
- geração de um conjunto de imagens de ambientes externos dinâmicos com mais de 3000 imagens panorâmicas, que poderão ser disponibilizadas para a comunidade acadêmica.

1.5 Estrutura da Dissertação

No presente capítulo foi apresentado o contexto, um breve estado da arte e o objetivo do trabalho. No Capítulo 2 é apresentado o embasamento teórico sobre o método SURF e discutida a motivação para a escolha deste como extrator de características. Na sequência, o Capítulo 3 apresenta o modelo desenvolvido para caracterização de cenas em ambientes externos; e mostra como aplicá-lo ao problema da localização robótica autônoma *online* em ambientes externos dinâmicos, e posteriormente à identificação de objetos.

No Capítulo 4 são apresentadas as métricas utilizadas para avaliar o desempenho do modelo proposto. No Capítulo 5 são apresentados os resultados das duas aplicações do modelo desenvolvido, e no Capítulo 6 são apresentadas as métricas utilizadas para medição de tempo, são discutidas as características temporais do modelo proposto e são apresentados os tempos de resposta dos experimentos realizados. Por fim, o Capítulo 7 apresenta as conclusões e os trabalhos futuros que podem ser desenvolvidos a partir dos resultados obtidos no presente trabalho.

Parte II

Fundamentação Teórica

2 Extração de Características

De acordo com o objetivo de desenvolver um método de localização visual autônomo *online* em ambientes externos dinâmicos, sem considerar informações *a priori*, são apresentados os problemas envolvidos na definição de um modelo para caracterização de cenas em ambientes externos dinâmicos. Primeiramente são apresentadas as características de imagens naturais, que estão fortemente presentes em ambientes externos; e em seguida é feita uma discussão sobre o método SURF (BAY *et al.*, 2008), que foi utilizado para extrair as características visuais do ambiente.

2.1 Imagens Naturais

O desafio de trabalhar em ambientes externos está fortemente ligado ao tratamento de imagens naturais, uma vez que na maioria desses ambientes sua presença é significativa. Por exemplo, mesmo caminhando em um ambiente urbano podemos observar vegetação, plantas, pequenos animais e de forma predominante o céu. Logo, é preciso utilizar técnicas apropriadas para representar elementos naturais.

Com discutido em Côco (2007), a caracterização de imagens naturais pode ser feita por texturas naturais, uma vez que texturas estão presentes em quase todos os objetos. No entanto, observa-se que cada autor apresenta a sua definição de textura, pois não existe uma única formalização precisa e nem uma aproximação matemática formal para sua caracterização.

Além disso, a grande variedade de propriedades que caracterizam texturas, dificulta a criação de uma única técnica de processamento para tratar imagens naturais (CôCO, 2007). Porém, um ponto comum à maioria das caracterizações de texturas é sua relação com a distância do observador à cena. Ou seja, quando uma região da imagem apresenta uma pequena variação tonal, a propriedade predominante é o tom dos pixels; por outro lado, quando esta variação é grande a propriedade predominante é a textura.

Contudo, as texturas não são as únicas ferramentas para modelar ambientes externos. A relação entre os espectros de potência em cenas de imagens naturais e a distância do observador à cena foi apresentada por Torralba e Oliva (2003). Neste trabalho é explicitado que imagens naturais apresentam assinaturas espectrais¹ diferentes de imagens artificiais, e ambas as assinaturas são deformadas com relação à distância do observador à cena.

Pode-se concluir dos trabalhos de Torralba e Oliva (2003) e Côco (2007) que imagens naturais apresentam fortes componentes de fractais estatísticos que se preservam

¹ Curvas de nível do espectro de potência da imagem.

nas bordas de elementos naturais. Logo, percebe-se que a escolha de um método extrator de características deve levar em consideração a capacidade de discriminação entre elementos da imagem, tanto naturais quanto artificiais, de forma invariante às condições de observação da cena.

Assim, o método SURF (BAY *et al.*, 2008) mostrou-se adequado para extrair informações de ambientes externos, visto que ele trabalha com as respostas da imagem a filtros digitais de bordas e assim consegue representar as informações dessas regiões. Este método consegue identificar pontos de interesse tanto em elementos naturais quanto artificiais, levando em consideração os parâmetros de variação de iluminação, rotação e escala, como será apresentado na Seção 2.2.

2.2 Método SURF

Ambientes externos são altamente dinâmicos, pois apresentam características como: ausência de marcos artificiais pré-definidos, desníveis no terreno, variação de iluminação e variabilidade de objetos no ambiente (SPERO, 2004). Dessa forma, métodos de localização aplicados a tais ambientes devem ser robustos o suficiente para considerar o seu caráter dinâmico. Por isso, optou-se por utilizar neste trabalho o método SURF como base para realização do processo de localização, uma vez que ele proporciona a obtenção de correspondências entre pontos de interesse presentes em imagens obtidas sob diferentes condições de observação.

Apesar de existirem diversos métodos de Visão Computacional que conseguem achar correspondências em imagens obtidas sob diferentes condições de observação, pode-se observar uma crescente utilização do método SURF, dado seu alto desempenho com relação à obtenção de correspondências (*matchings*) entre pontos de interesse, e o seu baixo tempo computacional se comparado a outros métodos semelhantes. O desempenho desse método já foi comprovado em alguns trabalhos, tais como (GAUGLITZ; HÖLLERER; TURK, 2011), que compara o desempenho de diversos métodos aplicados a um processo de rastreamento (*tracking*); e no próprio artigo de SURF (BAY *et al.*, 2008).

Assim, é preciso entender quais são os passos necessários para determinação de correspondências entre pontos de interesse, que segundo Bay *et al.* (2008) pode ser dividida em três passos principais:

1. detector de pontos de interesse (*detector*): pontos de interesse são selecionados em áreas distintas na imagem, tais como *corners*, *blobs*², e *T-junctions*, sendo sua

² Região de uma imagem digital na qual algumas propriedades são constantes ou variam dentro de uma faixa predeterminada. Todos os pontos em um *blob* podem ser considerados, de alguma forma, similares uns aos outros.

- propriedade mais valiosa a repetibilidade, ou seja, a capacidade de encontrar de forma confiável o mesmo ponto de interesse sob diferentes condições de observação;
2. descritor de pontos de interesse (*descriptor*): a vizinhança de cada ponto de interesse é representada por um vetor de características, e tal descritor deve ser distintivo e ao mesmo tempo robusto a ruído, a erros de detecção, e a deformações geométricas e fotométricas;
 3. detector de correspondências (*matching*): os vetores que caracterizam os pontos de interesse de imagens distintas são confrontados para se determinar as correspondências, sendo esse processo, muitas vezes, baseado na determinação de distâncias entre vetores, por exemplo a distância de Mahalanobis e a distância Euclidiana. Nesse processo, a dimensão do vetor de características apresenta impacto direto sobre o tempo de computação, logo uma baixa quantidade de dimensões é desejável.

2.2.1 Detector de Ponto de Interesse

A seleção de pontos de interesse é efetuada através da utilização de uma aproximação para Matriz Hessiana (*Hessian-Matrix*). O cálculo dessas aproximações é feito com a utilização de imagens integrais (VIOLA; JONES, 2001), uma vez que tal abordagem reduz o tempo computacional drasticamente.

A imagem integral $I_{\Sigma}(x, y)$ é formada através de uma transformação feita sobre a imagem original $I(x, y)$. Essa transformação substitui o valor de cada pixel da imagem pela soma de todos os pixels dentro da região formada pelo retângulo definido pelas coordenadas do pixel e da origem da imagem, como pode ser visto na equação a seguir

$$I_{\Sigma}(x, y) = \sum_{i=0}^{i < x} \sum_{j=0}^{j < y} I(x, y). \quad (2.1)$$

A vantagem na utilização de imagens integrais está no fato de que uma vez geradas, são necessárias apenas três somas para calcular a soma da intensidade de áreas retangulares como as da Figura 4. Assim, o tempo de computação dessa soma é independente do tamanho da região, o que é extremamente importante uma vez que o método SURF usa filtros de tamanho elevado.

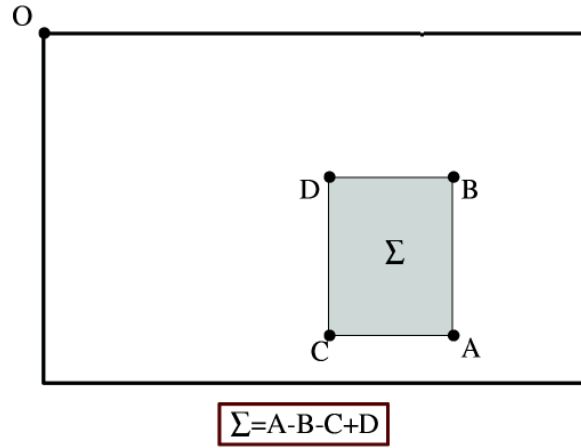


Figura 4 – Cálculo da soma de intensidade de pixels em uma região retangular usando imagens integrais. Fonte: [Bay et al. \(2008\)](#).

O detector utilizado no método SURF é baseado na Matriz Hessiana por causa do seu bom desempenho e precisão. Este procura por áreas da imagem que apresentem estrutura de *blob* em locais onde o determinante da Matriz Hessiana é máximo. Dado um ponto $X = (x, y)$ em uma imagem I , a Matriz Hessiana desse ponto $H(X; \sigma)$ à escala σ é definida por

$$H(X; \sigma) = \begin{bmatrix} L_{xx}(X, \sigma) & L_{xy}(X, \sigma) \\ L_{xy}(X, \sigma) & L_{yy}(X, \sigma) \end{bmatrix}, \quad (2.2)$$

onde $L_{xx}(X, \sigma)$ é a convolução da derivada de segunda ordem da Gaussiana $\frac{\delta^2}{\delta x^2} g(\sigma)$ com a imagem I no ponto X , e analogamente para $L_{xy}(X, \sigma)$ e $L_{yy}(X, \sigma)$. Gaussianas são muito utilizadas para análise de espaço de escala ([KOENDERINK, 1984](#); [LINDBERG, 1990](#)), porém para se alcançar um bom desempenho prático é necessário discretizá-las e limitá-las. Para tal, SURF utiliza aproximações para as derivadas de segunda ordem da Gaussiana como mostrado na Figura 5.

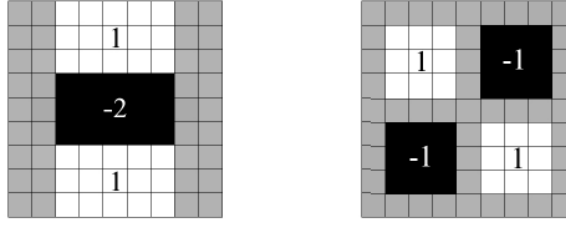


Figura 5 – Da esquerda para direita tem-se a aproximação utilizada no método SURF respectivamente para derivada parcial de segunda ordem da Gaussiana na direção y (D_{yy}) e xy (D_{xy}). As regiões em cinza apresentam valores iguais a zero. Fonte: Bay *et al.* (2008).

Os filtros 9x9 mostrados na Figura 5 são aproximações para as derivadas de uma Gaussiana com $\sigma = 1, 2$, que representam a menor escala para calcular as respostas de um *blob*. A convolução das aproximações da derivada da Gaussiana (D_{xx} , D_{yy} e D_{xy}) com o *blob* são representadas por L_{xx} , L_{yy} e L_{xy} . O parâmetro σ é utilizado como peso para calcular o determinante da aproximação para Matriz Hessiana, como mostrado a seguir

$$\det(H_{aproximada}) = L_{xx}(X, \sigma)L_{yy}(X, \sigma) - (\sigma L_{xy}(X, \sigma))^2. \quad (2.3)$$

Analisando a Equação 2.3, pode-se concluir que o determinante da Matriz Hessiana muda conforme a escala utilizada para calcular a resposta do *blob*. Porém, na prática, o método SURF mantém esse fator constante e normaliza a resposta do *blob* com relação à escala do filtro.

A aproximação para o determinante da Matriz Hessiana representa a resposta do *blob* presente na posição X da imagem. Para cada *blob* são calculadas respostas em diferentes escalas, pois a busca por correspondências entre pontos de interesse normalmente é aplicada a imagens que representam cenas em escalas diferentes, como exemplificado na Figura 6. Para isso, a teoria de Espaço de Escalas (LINDBERG, 1990), representada através de uma pirâmide como a da Figura 7, é utilizada juntamente com imagens integrais para calcular a resposta dos *blobs* em diferentes escalas. O processo consiste em suavizar as imagens repetidamente (camada a camada da pirâmide) através de Gaussianas (filtros) e então subamostrar o resultado para atingir um nível mais alto na pirâmide, ou seja, a imagem filtrada a uma escala maior.



Figura 6 – A cabana dentro do círculo laranja representa o objeto de interesse. À esquerda temos a cabana em uma escala menor dentro da cena e à direita, após uma operação de *zoom in*, temos a cabana em uma escala maior.

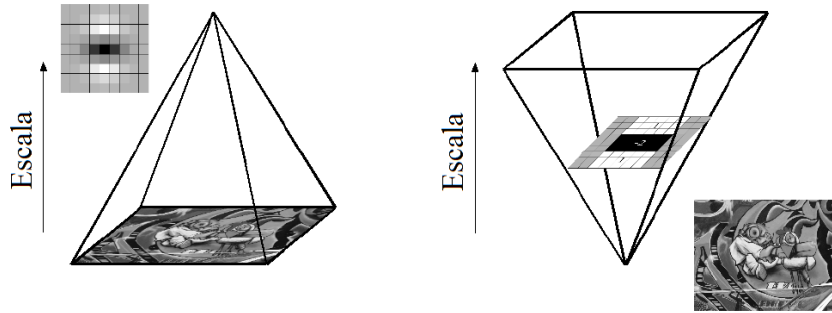


Figura 7 – Aplicação do espaço de escalas. Ao invés de reduzir o tamanho da imagem (esquerda) as imagens integrais são utilizadas para aumentar a escala do filtro (direita) a um custo computacional constante. Fonte: [Bay et al. \(2008\)](#).

O espaço de escalas começa com um filtro de dimensão 9×9 , que calcula a resposta dos *blobs* da imagem para a menor escala considerada. Na sequência, são aplicados à imagem os filtros de dimensão 15×15 , 21×21 , e 27×27 , que conseguem alcançar uma mudança de escala superior a 2. No entanto, é necessário que uma redução 3D não-máxima (*3D non-maximum suppression*)³ seja aplicada tanto espacialmente quanto ao longo das escalas vizinhas. Consequentemente, o primeiro e o último mapa Hessiano de resposta, ou seja, resposta dos *blobs* da imagem aos filtros na primeira e na última escala, não poderão ser considerados máximos pois esses são tomados apenas como referência.

A partir deste ponto será utilizado símbolo e para determinar o valor real da escala, que apesar de depender diretamente do desvio padrão σ das derivadas das Gaussianas utilizadas para calcular as respostas, não é definida apenas por este valor.

³ A imagem é varrida ao longo da direção de seu gradiente, e se um pixel não faz parte dos máximos locais em relação ao módulo de seu gradiente, seu valor é alterado para zero. O efeito dessa operação é excluir toda informação da imagem que não faz parte de seu máximo local.

Portanto, após aplicar o método de interpolação proposto por [Brown e Lowe \(2002\)](#), que considera a escala e o tamanho da imagem para se encontrar o máximo das respostas, tem-se que a menor escala considerada é $e = 1,6 = 1,2 \times 12/9$, que corresponde a um filtro de tamanho 12×12 , e a maior é $e = 3,2 = 1,2 \times 24/9$, que corresponde a um filtro de tamanho 24×24 . Dessa forma, o algoritmo de detecção desenvolvido em ([BAY et al., 2008](#)) (*Fast-Hessian detector*) consegue identificar pontos de interesse em várias escalas, como por exemplo os pontos no campo de girassóis da Figura 8, no qual os círculos brancos representam a região do ponto de interesse e sua respectiva escala.

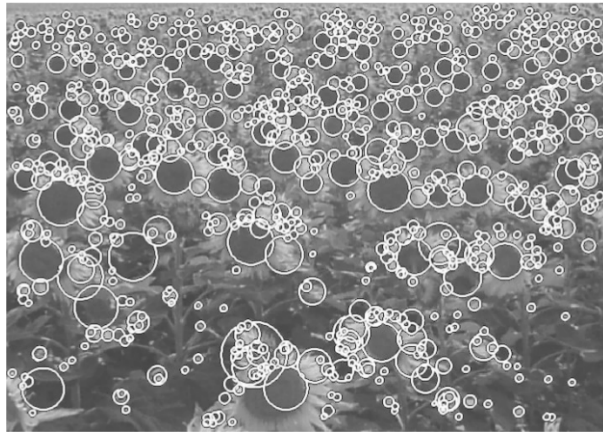


Figura 8 – Pontos de interesse detectados em um campo de girassóis pelo método detector *Fast-Hessian detector*. Fonte: [Bay et al. \(2008\)](#).

Logo, com o uso de aproximações para derivadas de segunda ordem da Gaussiana (*box filters*) e de imagens integrais, o método SURF não requer a aplicação iterativa de um filtro à imagem de saída do nível anterior da pirâmide (Figura 7) ([BAY et al., 2008](#)). Assim, para se determinar a resposta de um *blob* é aplicado um filtro de qualquer tamanho diretamente à imagem original.

A aplicação do filtro 9×9 da Figura 5 é considerado como a camada inicial da pirâmide, que apresenta a menor escala (aproximação da derivada de segunda ordem da Gaussiana com desvio padrão σ igual a 1,2). Para obtenção das próximas camadas da pirâmide são aplicados filtros com tamanhos gradualmente maiores. Assim, levando em conta a natureza discreta de imagens integrais e a estrutura dos filtros propostos, o método SURF é computacionalmente viável para detecção de pontos de interesse em imagens digitais.

2.2.2 Descritor de Ponto de Interesse e Determinação de Correspondências

O descritor desenvolvido por ([BAY et al., 2008](#)) é baseado na distribuição da intensidade dos pixels da vizinhança em torno do ponto de interesse, similar ao descritor desenvolvido no método SIFT ([LOWE, 1999](#)). Esse descritor é baseado na distribuição de

primeira ordem das respostas *Haar Wavelet* nas direções x e y ao invés do gradiente, utiliza imagens integrais, e apresenta dimensão igual a 64. Todos esses componentes reduzem o tempo de computação das características e principalmente do algoritmo que determina as correspondências entre características. Além disso, foi comprovado em [Bay et al. \(2008\)](#) que esse descritor, apesar de ser otimizado do ponto de vista de tempo computacional, também se mostra robusto.

Para calcular os descritores e em seguida determinar correspondências, são realizados os seguintes passos:

- Passo 1: é determinada a orientação de uma região circular em torno do ponto de interesse;
- Passo 2: é construída uma região quadrada alinhada à orientação da região determinada no passo 1, para então extrair o descritor SURF;
- Passo 3: um algoritmo determina correspondências entre características (ponto de interesse + descritor) extraídas de duas imagens diferentes, uma vez que todos os pontos tiveram seus descritores calculados.

2.2.2.1 Passo 1: determinação da orientação

Como é desejado um algoritmo de determinação de correspondências que seja invariante à rotação da imagem, é necessário identificar a orientação do ponto de interesse, ou seja, a orientação do descritor desse ponto, para que posteriormente essa informação seja considerada na hora de determinar as correspondências. Para tal, são calculadas as respostas *Haar Wavelet* nas direções x e y dentro de uma vizinhança circular de raio $6e$ em torno do ponto de interesse, onde e é a escala na qual o ponto foi detectado.

Vale ressaltar que a escolha dos pontos da vizinhança que serão considerados no descritor é definida por um passo de amostragem igual a e , e o tamanho dos filtros *Haar Wavelet* também é dependente da escala, e definido com tamanho lateral igual a $4e$. Segundo [Bay et al. \(2008\)](#), com a utilização desses filtros, Figura 9, são necessárias apenas seis operações para calcular a resposta na direção x ou y em qualquer escala, dado que são utilizadas imagens integrais que aceleram o processo.

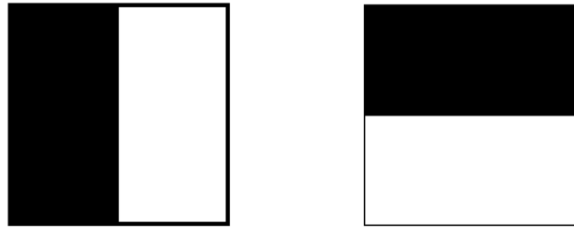


Figura 9 – Filtros *Haar Wavelet* para calcular as respostas na direção x (esquerda) e na direção y (direita). As partes pretas do filtro apresentam peso -1 e as partes brancas $+1$. Fonte: Bay *et al.* (2008).

Depois de calcular as respostas *Haar Wavelet* da vizinhança e suavizá-las com uma Gaussiana ($\sigma = 2e$) centrada do ponto de interesse, estas respostas são representadas como pontos em um espaço bidimensional, no qual o valor da resposta na direção x corresponde à abscissa do espaço e o valor da resposta na direção y corresponde à ordenada. A orientação dominante de uma região é estimada através da soma de todas as respostas dentro de uma janela deslizante de orientação de tamanho $\pi/3$ (setor circular na cor cinza), como apresentado na Figura 10. A resposta horizontal e vertical são compostas para definir o vetor de orientação local da região, e o maior vetor dentre as janelas da vizinhança define a orientação do ponto de interesse.

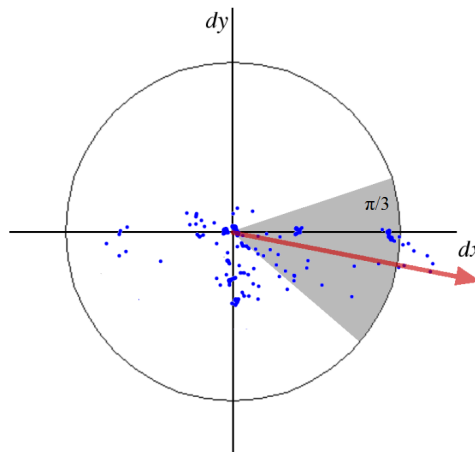


Figura 10 – Determinação da orientação de um ponto de interesse: uma janela deslizante de orientação de tamanho $\pi/3$ determina a orientação dominante das respostas *Haar Wavelet* suavizada por uma Gaussiana, definidas para cada ponto de amostragem dentro de uma vizinhança circular em torno do ponto de interesse. Fonte: Bay *et al.* (2008).

É importante deixar claro que não são todas as aplicações que necessitam de um método para determinação de correspondências que seja invariante à rotação, como em

situações que só existe mudança na escala da imagem. Logo, nessas aplicações pode-se utilizar a versão do método SURF chamada *Upright Surf* (U-SURF), que é mais rápida pois não considera rotação dos pontos e ainda assim consegue se manter robusta para inclinações na faixa de ± 15 graus.

2.2.2.2 Passo 2: extração de informação

Para definir as regiões orientadas que servirão de entrada para o cálculo dos descritores, é construído um quadrado de tamanho $20e$, centrado no ponto de interesse e orientado segundo os passos descritos anteriormente. Na Figura 11 pode-se observar um exemplo dessas regiões.



Figura 11 – Imagem que mostra o tamanho da janela das regiões orientadas em diferentes escalas. Fonte: [Bay et al. \(2008\)](#).

Essa região é dividida regularmente em 16 sub-regiões quadradas, através de uma divisão 4×4 da região original. Então, para cada uma dessas sub-regiões são calculadas respostas *Haar Wavelet* em pontos definidos por uma divisão 5×5 , que determinam um espaçamento regular entre os pontos da sub-região. Considere dx como sendo a resposta *Haar Wavelet* na direção horizontal e dy a resposta na direção vertical para filtros de tamanho $2e$. Nesse caso, os termos “horizontal” e “vertical” são definidos relativamente à orientação do ponto de interesse, como pode ser visto na Figura 12. Para aumentar a robustez à deformações geométricas e erros de localização, as respostas dx e dy são suavizadas por uma Gaussiana ($\sigma = 3,3e$) centrada no ponto de interesse. Então, as respostas dx e dy de cada sub-região são somadas para formar informações do vetor descritor de uma característica.

Além desse conjunto de informações, são consideradas informações sobre a polarização das mudanças de intensidade ao longo da região. Assim, é calculada a soma dos valores absolutos das respostas dx e dy ($|dx|$ e $|dy|$). Consequentemente, cada sub-região apresenta

um vetor descritor v de quatro dimensões $v = [dx, dy, |dx|, |dy|]$. Concatenando esse vetor v para todas as 16 (4x4) sub-regiões, resultará em vetor descritor para o ponto de interesse de tamanho 64.

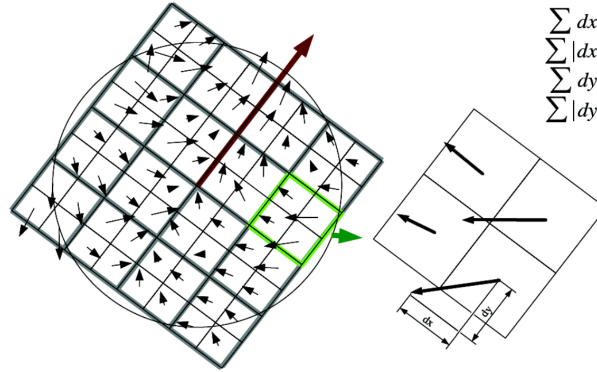


Figura 12 – Para construir o descritor, é colocado sobre o ponto de interesse um quadrado orientado dividido em 4x4 sub-regiões quadradas (esquerda). Nas divisões 2x2 (direita) das sub-regiões 4x4 são calculadas as respostas *Haar Wavelet*, dx e dy , relativamente à orientação do ponto. Para montar o descritor v , são computadas as somas $\sum dx$, $\sum |dx|$, $\sum dy$, e $\sum |dy|$ para cada sub-região 4x4. Fonte: Bay *et al.* (2008).

Também foi desenvolvido por Bay *et al.* (2008) um descritor similar que apresenta uma vizinhança de tamanho 128 chamado SURF-128. Este descritor apresenta as mesmas somas descritas anteriormente, porém com uma resolução maior. As somas de dx e $|dx|$ são computadas separadamente para $dy < 0$ e $dy \geq 0$. De forma análoga, as somas de dy e $|dy|$ são calculadas em duas partes dependendo do sinal de dx , o que proporciona um descritor de tamanho duplicado. Consequentemente, o descritor SURF-128 mostra-se mais distintivo e não perde muita velocidade de processamento, apesar de ficar mais lento, se comparado ao descritor SURF-64. Por isso, e por outros motivos expostos na Seção 2.2.3 optou-se neste trabalho pela utilização da versão estendida do método SURF, o SURF-128.

2.2.2.3 Passo 3: determinação de correspondências

Depois de detectar pontos de interesse e calcular os seus descritores já é possível procurar por correspondências entre características. Assim, para que essa busca por correspondências não inviabilize o método, SURF utiliza uma estrutura de indexação rápida baseada em árvores multidimensionais chamada *k-dimensional trees* (*k-d trees*). Para que essa estrutura seja utilizada é necessária a definição de um parâmetro para montar a árvore em suas várias dimensões, que no algoritmo de *matching* SURF é definido pelo sinal do traço da Matriz Hessiana do ponto de interesse.

Como normalmente pontos de interesse são detectados em regiões com estruturas de *blob*, o sinal do Laplaciano propicia a distinção entre *blobs* brilhantes sob fundos escuros e *blobs* escuros sob fundos brilhantes. O mais importante é que esse parâmetro está disponível desde a etapa de detecção, sem que se tenha um custo computacional extra. Então, antes de efetuar a comparação entre as características, o algoritmo de *matching* checa se as características a serem comparadas apresentam o mesmo tipo de contraste, como exemplificado na Figura 13. Por isso, essa simples informação consegue acelerar o algoritmo de *matching*, sem reduzir seu desempenho.

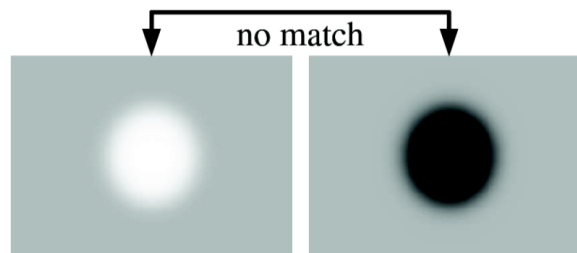


Figura 13 – Se o contraste entre dois pontos de interesse é diferente, esses candidatos a correspondentes são descartados. Fonte: Bay *et al.* (2008).

2.2.3 Avaliação do método SURF

Para avaliar o desempenho do método SURF os próprios autores Bay *et al.* (2008) apresentaram alguns testes utilizando diferentes critérios. Para tal, foi utilizado um conjunto de imagens de cenas estruturadas, que apresentavam textura e que foram geradas através de diferentes transformações geométricas e fotométricas tais como: mudança de ponto de observação, rotação, zoom, borrão, mudança do nível de luminosidade e compressão *Joint Photographic Experts Group* (JPEG).

Inicialmente foi analisado o desempenho apenas do detector de pontos de interesse (*Fast-Hessian Detector*), considerando duas versões desse detector em relação ao filtro Gaussiano que define a menor escala de referência: o FH-9, que representa o detector com filtro inicial de tamanho 9x9; e o FH-15, que apresenta filtro inicial de tamanho 15x15. Essas duas versões do detector foram avaliadas segundo o critério de repetibilidade ((SCHMID; MOHR; BAUCKHAGE, 2000)) e comparadas aos detectores clássicos: *Difference of Gaussians* (DoG) (LOWE, 2004), *Harris-Laplace* e *Hessian-Laplace* (MIKOLAJCZYK; SCHMID, 2004), como mostram os gráficos da Figura 14.

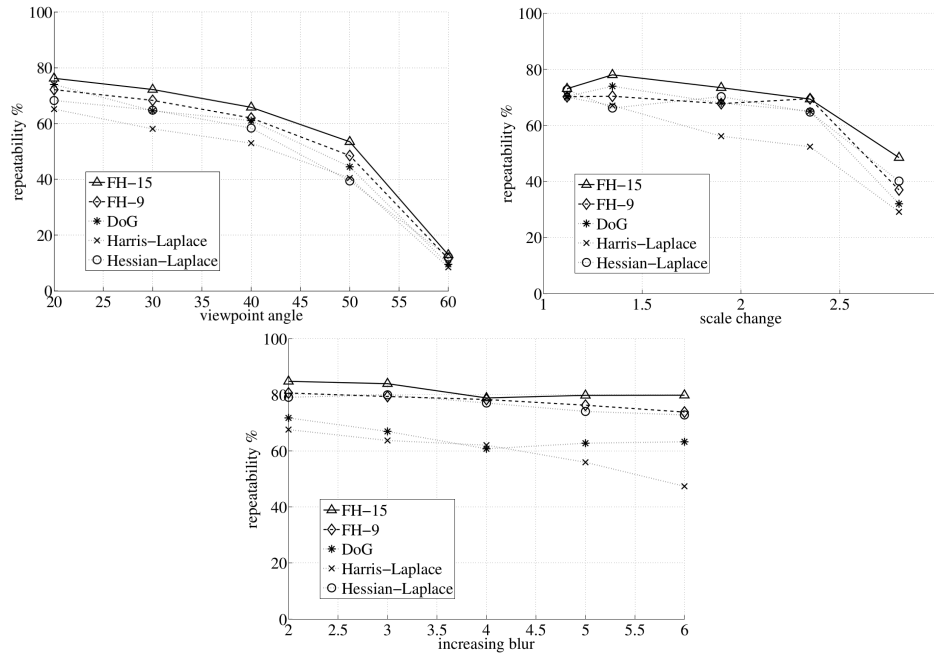


Figura 14 – Avaliação de desempenho das versões do detector SURF, FH-9 e FH-15, através do critério de avaliação de repetibilidade considerando respectivamente as seguintes características das imagens: mudança de ponto de vista (acima à esquerda), mudança de escala (acima à direita) e adição de níveis de borrão (abaixo). Fonte: Bay *et al.* (2008).

Segundo (BAY *et al.*, 2008) o detector FH-9 é 5 vezes mais rápido que o DoG e 10 vezes mais rápido que *Hessian-Laplace*. Já, o detector FH-15 é mais de 3 vezes mais rápido que o DoG e mais de 4 vezes que o *Hessian-Laplace*. Com relação ao critério de repetibilidade, tanto o FH-9 quanto o FH-15 são pelo menos comparáveis aos outros métodos utilizados como referência, mas na maioria das situações os superam como pode ser visto na Figura 14.

Posteriormente, foram analisadas as diferentes variações do descritor SURF e os seus efeitos através do método de avaliação *recall/precision* (DAVIS; GOADRIC, 2006). Para alterar os descritores da versão padrão foi modificada a quantidade de subdivisões feitas na região quadrada da Figura 12, o que impacta diretamente na velocidade do algoritmo de *matching*, pois descritores maiores implicam em um maior número de comparações.

A Figura 15 mostra a avaliação efetuada tanto para os descritores padrão, de cujas sub-regiões definem dx , dy , $|dx|$, $|dy|$, quanto para os descritores estendidos, que apresentam o dobro do tamanho dos descritores padrão pois dx e $|dx|$ são calculados em relação ao sinal de dy e vice-versa. Considerando divisões com os seguintes tamanhos: 3x3 (SURF-36), 4x4 (SURF-64), 5x5 (SURF-100) e 6x6 (SURF-144); e suas respectivas versões estendidas, pode-se observar que a versão SURF-64 e sua versão estendida SURF-128 apresentam um melhor desempenho que as outras testadas, visto que em curvas de

avaliação do tipo *recall/precision*, quanto maior for a área debaixo da curva, e/ou mais próximo do canto superior direito do quadrado formado pelos pontos (0,0) e (1,1) esta curva estiver, melhor é seu desempenho.

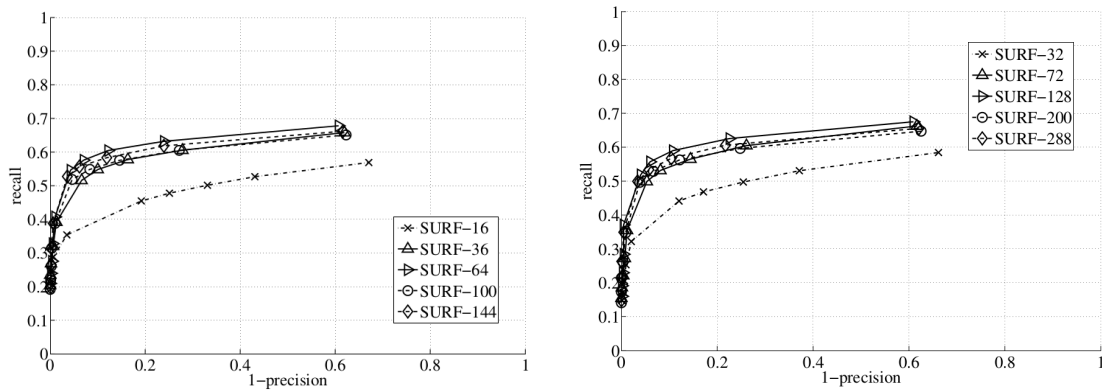


Figura 15 – Curvas *recall/precision* geradas pelo algoritmo de *matching* baseado no vizinho mais próximo (*Nearest Neighbor*), considerando variações dos descritores SURF tanto para as versões padrão (à esquerda), quanto para suas versões estendidas (à direita). Fonte: Bay *et al.* (2008).

Uma vez determinada a melhor versão do descritor SURF, foi avaliado o seu desempenho aplicado ao problema de reconhecimento de objetos. Nesse experimento foi utilizado uma implementação disponível para o classificador *Naive Bayes* através do modelo de objetos *two bag-of-words* (FERGUS; PERONA; ZISSERMAN, 2003). O objetivo do teste é: dados uma imagem e um objeto, determinar se esse objeto pertence à imagem ou não. Os testes foram baseados em 400 imagens das quais 50% foram utilizadas para treinamento do classificador e o restante para efetuar os testes. Como o objetivo era avaliar apenas o desempenho do descritor frente a outros descritores já consagrados, foi utilizada a geração automática de pontos de interesse fornecida pelo *framework* (FERGUS; PERONA; ZISSERMAN, 2003), que baseia suas escolhas em pontos de borda; e também a geração de pontos de interesse determinados pelo detector SURF.

Pode-se observar na Figura 16 as curvas de *Receiver Operating Characteristics* (ROC) (FAWCETT, 2006), que apresentam o desempenho do descritor SURF-128 frente aos descritores SIFT e *Gradient Location and Orientation Histogram* (GLOH). Tanto para a descrição de pontos de interesse definidos pelo detector SURF quanto pelo *framework*, o descritor SURF-128 consegue superar os demais, visto que a conclusão sobre curvas ROC é similar à conclusão sobre curvas *recall/precision*⁴, pois quanto maior for a área debaixo da

⁴ Curvas ROC e *recall/precision* apresentam mapeamento entre si e podem ser comparadas de forma direta (DAVIS; GOADRIC, 2006). No entanto, os índices utilizados são diferentes e os limiares de classificação, nos extremos, avaliados nas curvas ROC (1 e 0) são diferentes dos avaliados em curvas *recall/precision*.

curva, e/ou mais próximo do canto superior esquerdo do quadrado formado pelos pontos (0,0) e (1,1) estiver a curva, melhor é seu desempenho.

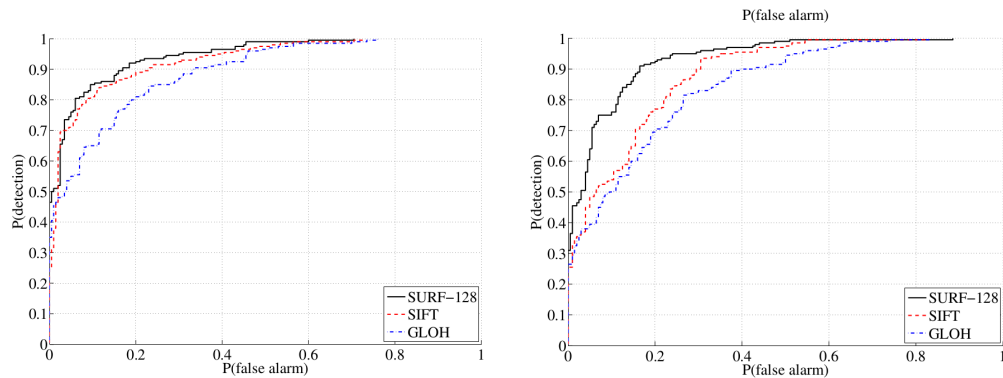


Figura 16 – Comparação entre curvas ROC de classificadores *Naive Bayes* utilizando o modelo para representação de objetos *bag-of-words*, aplicado a diferentes métodos descritores. Os pontos de interesse utilizados pelos métodos descritores foram gerados de duas formas: através da determinação aleatória de píxeis de borda (à esquerda), e pelo detector SURF (à direita). Fonte: [Bay et al. \(2008\)](#).

Além de todos os resultados positivos apresentados pelos próprios autores do método SURF, é possível encontrar alguns trabalhos que validam o desempenho desse método. Em ([CUMMINS; NEWMAN, 2010](#)) e ([MITSUHASHI; KURODA, 2011](#)) são apresentadas aplicações que utilizam o método SURF como etapas de extração e descrição de características visuais, nos quais os resultados apresentados mostram o bom desempenho desse método.

Por fim, em ([MITSUHASHI; KURODA, 2011](#)) é feita uma comparação criteriosa e detalhada sobre 30 combinações de métodos detectores/descritores aplicados ao problema de rastreamento visual, no qual é validado mais uma vez a boa relação custo/benefício do método detector/descritor SURF, já constatados em ([BAY et al., 2008](#)). Neste trabalho foram efetuados testes separados sobre os detectores e descritores dos métodos considerados e por fim foi avaliada a influência destes no processo de rastreamento visual. Nesse trabalho não foi destacado um único método, mas foram apresentados vários resultados destacáveis principalmente do detector SURF, *Fast Hessian Detector*.

Dessa forma, após um estudo detalhado do método SURF e a confirmação de que a relação custo x benefício deste mostra-se muito boa se comparada a outros métodos consagrados, optou-se pela utilização deste como método para extração de características de ambientes externos a ser utilizado pelo modelo proposto neste trabalho e discutido no Capítulo 3.

Parte III

Modelagem

3 Caracterização de Cenas em Ambientes Externos

O método SURF consegue extrair características de imagens e determinar correspondências entre elas de forma robusta, além de apresentar uma boa relação entre custo computacional e precisão. Assim, ele pode ser utilizado para criação do modelo para caracterização de cenas proposto nesse trabalho.

Contudo, apesar de sua robustez e invariância o método SURF ainda é susceptível a falhas na determinação de correspondências, como pode-se observar na Figura 17. À esquerda são apresentadas duas imagens da mesma cena sobre condições de observação diferentes, sob as quais as correspondências são determinadas precisamente. Por outro lado, à direita, são apresentadas duas imagens de cenas diferentes, sob as quais são determinadas correspondências erradas, mostrando assim a possibilidade de falha do método SURF.

Dessa forma, o modelo para caracterização de cenas deve levar isso em conta para que seja possível manter a robustez do extrator de características. Por isso, foi utilizada uma abordagem probabilística para tal modelo.



Figura 17 – À esquerda correspondências corretas e à direita correspondências erradas.

3.1 Caracterização de Cenas Aplicada ao Problema de Localização: O Modelo Proposto

Analizando o problema de localização apresentado no Capítulo 1, é proposto aqui uma abordagem probabilística para modelá-lo. Ao invés de implementar uma localização determinística (ex: “a imagem Im_x pertence ao lugar L_5 ”) nesse trabalho gera-se a probabilidade da imagem pertencer a cada local do mapa para que seja possível considerar incertezas no processo de localização, aumentando assim a robustez do método de localização utilizado sobre o modelo.

A robustez é aumentada pois modelos probabilísticos representam as incertezas do mundo real de forma explícita, e assim podem ponderar vários lugares ao invés de indicar uma única localização no mapa. Logo, essa abordagem permite que os algoritmos de tomada de decisão tratem ambiguidades e graus de certeza em uma boa forma matemática (THRUN; BURGARD; FOX, 2005).

No entanto, não será considerada nenhuma informação *a priori* sobre a localização do robô, como apresentado no modelo bayesiano da Figura 1. Pois, definindo um método de localização que não considere informações de tempos passados, é possível expandi-lo facilmente para um modelo de localização bayesiano completo, através do armazenamento das informações ao longo do percurso do robô. Tal abordagem é sugerida como trabalho futuro, com o objetivo de aumentar a taxa de acerto e o desempenho do método de localização.

3.1.1 Caracterização de Lugares

O modelo probabilístico deve possibilitar a localização de um grupo de imagens que caracterize determinado lugar. Para isso, é necessário calcular as probabilidades de um conjunto de características, extraído do grupo de imagens¹ de caracterização do lugar em questão, pertencer aos demais lugares do ambiente que constituirão um mapa. Dessa forma, a motivação para o desenvolvimento do modelo aqui proposto é possibilitar o cálculo de tal probabilidade utilizando somente a aparência do lugar.

Para isso, cada lugar do mapa é representado por uma função massa de probabilidade (PMF) estimada por um histograma de características. Como esse trabalho foi desenvolvido sobre o método SURF, as características utilizadas na PMF serão as extraídas por este método.

No entanto, vale ressaltar que pode-se utilizar o mesmo conceito sobre características diferentes, tanto baseadas em aparência quanto métricas, tais como SIFT, medidas de ultrassom, ou até operadores de cor. Porém, a escolha de tais características deve levar em

¹ O tamanho desse grupo depende do campo de visão (FOV - *Field Of View*) da câmera utilizada.

conta os fatores tempo computacional e invariabilidade aos parâmetros de observação da cena, para que os resultados obtidos sejam satisfatórios. Todavia, conforme já mencionado, este trabalho se limitou à utilização de características visuais SURF.

Para calcular a PMF de um lugar será necessário definir a variável aleatória L . Considere o experimento E , que representa a determinação da localização de um grupo de imagens quaisquer, e ξ um possível resultado desse experimento. O espaço amostral de E será definido pelos lugares do ambiente mapeado, por exemplo, para o mapa da Figura 18, os únicos possíveis valores para ξ são: $L_0, L_1, L_2, L_3, L_4, L_5, L_6, L_7, L_8$ e NE (lugar não mapeado). Então, considerando uma função identidade $f(\xi) = \xi$, uma realização da variável L é definida pela equação $L = f(\xi)$.

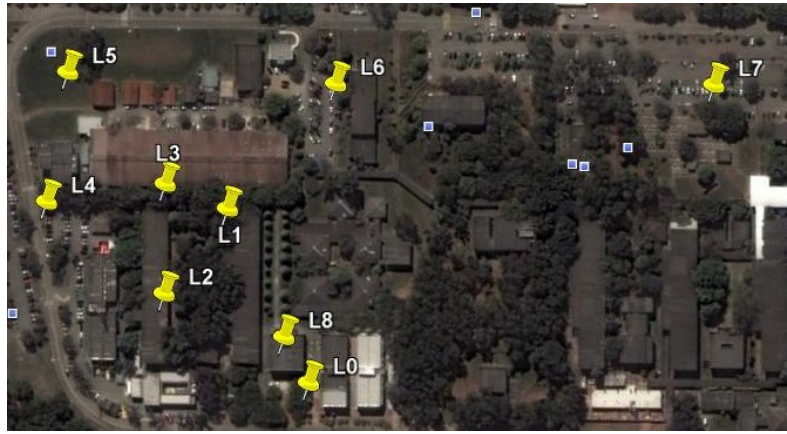


Figura 18 – Exemplo de ambiente externo e seus lugares L_i representados pelos marcadores amarelos. Observe que L_i é a i -ésima realização da variável aleatória L .

A definição da variável aleatória L é necessária para formalizar o cálculo da probabilidade Pr de um conjunto de características pertencer a um lugar do mapa, pois de acordo com a definição de PMF apresentada por [Miller e Childers \(2004\)](#), temos que $PMF_L(L_i) = Pr(L = L_i)$. Porém, a determinação de PMF_L não é uma tarefa simples, pois essa função depende da postura do robô, a qual depende do comportamento do mesmo no ambiente.

Contudo, a localização L_i pode ser representada através da PMF de um conjunto de características s_i extraídas do grupo de imagens utilizado para caracterizar o lugar L_i , como pode ser observado em

$$PMF_L(L_i) = PMF_{s_i}. \quad (3.1)$$

A Equação 3.1 pode ser traduzida pela seguinte sentença “A PMF_L de um ambiente, avaliada em um lugar qualquer L_i , é igual à PMF de um conjunto de características s_i extraído deste lugar”. Dessa forma, a probabilidade da variável aleatória L , calculada

pelos histogramas de características dos lugares, servirá para determinar a localização do robô que capturou tal imagem ou, generalizando, a localização da câmera, que pode estar acoplada a um equipamento qualquer ou sendo utilizada por uma pessoa.

Porém, para calcular esta probabilidade é necessário definir as PMFs dos lugares do mapa, que são estimadas pelos histogramas dos índices das características SURF destes lugares. Como tal histograma é definido pelos índices e não pelas características, caracteriza-se uma redução de dimensão de 128 para 1 no espaço das características, que é alcançada pela utilização do algoritmo de determinação de correspondências SURF.

É importante ressaltar que as PMFs dos lugares também podem ser estimadas por outros métodos de inferência, tais como Janelas de Parzen e *K-Nearest Neighbors* (K-NN). Contudo, esses métodos aumentariam o tempo de cálculo das PMFs, visto que seria necessário ordenar as características antes de computar suas ocorrências nos volumes de domínio das PMFs.

Logo, com a utilização de histogramas para inferir as PMFs não é preciso ordenar as características, pois somente são computadas as ocorrências destas nos lugares. Então, para definir as PMFs são realizados os seguintes passos para cada lugar do mapa,

1. **é capturado um conjunto de imagens do lugar;**
2. **são extraídas as características SURF desse conjunto de imagens;**
3. **a ocorrência dessas características é computada sobre o conjunto de imagens, utilizando o algoritmo de determinação de correspondências;**
4. **através do cálculo de frequência relativa das características é obtido um histograma de características SURF para o conjunto de imagens do lugar;**
5. **o histograma de características é normalizado e utilizado como aproximação da PMF do lugar.**

Sabe-se que para determinar uma PMF de forma correta é necessário utilizar amostras que sejam representativas sob o ponto de vista da variável aleatória em questão. Dessa forma, é de fundamental importância a escolha do conjunto de imagens que dará origem às amostras de características presentes em um lugar, pois tal escolha pode influenciar muito a capacidade de discriminação do modelo de caracterização de cenas.

No entanto, tendo em vista a questão da utilização desse modelo em métodos de localização autônoma e *online*, levantada no Capítulo 1, define-se a seguinte restrição: durante processos de localização *online* tais como SLAM, não é possível capturar de forma *online* imagens que apresentem características do lugar sob diversas condições ambientais,

uma vez que o tempo de captura das imagens seria muito reduzido para não prejudicar o processo de navegação do robô.

Assim, a PMF gerada de forma *online* poderia não ser representativa e ocasionar falhas nos métodos de localização utilizados sobre o modelo. No entanto, para resolver esse problema, pode-se utilizar mapas contruídos de forma *offline*, e caso seja necessário a utilização de SLAM, o tempo gasto pelo método de forma *online* seria reduzido, pois só seria necessário atualizar as PMFs com novas características encontradas. Logo, este trabalho se restringiu a caracterizar as cenas de forma *offline*.

O processo de captura das imagens também é muito importante, pois pode favorecer ou prejudicar a extração das características de um lugar. Assim, será definido um conceito novo, aqui denominado de amostra completa: amostra definida pela captura de uma realização da variável aleatória em questão, que possa conter valores que representem por completo esta variável.

A determinação prática do que vem a ser uma amostra completa depende da cena a ser caracterizada e do campo de visão (FOV - *Field Of View*) da câmera utilizada para capturar as imagens. Por exemplo, considerando a caneca apresentada na Figura 19 são necessárias 8 imagens obtidas por uma câmera com FOV de aproximadamente 60°, para se obter informações de todos os pontos de vista externos da mesma. Para uma situação na qual deseja-se caracterizar também o interior e o fundo da caneca seriam necessárias mais duas imagens para compor uma amostra completa da cena.



Figura 19 – Exemplo de amostra completa para uma caneca.

No problema em questão, o objetivo é caracterizar lugares pertencentes a um ambiente externo. Logo, a definição de uma amostra completa para essa cena seria um conjunto de imagens que juntas totalizam um campo visual de 360°² do lugar. Assim, através da utilização de uma câmera grande angular, que apresenta FOV de aproximadamente 180°, foram capturadas 4 imagens para compor uma amostra completa da variável aleatória em questão. Optou-se por usar 4 imagens porque o FOV da câmera não é exatamente

² Nesse caso, desconsideram-se informações sobre céu e chão, uma vez que estas normalmente não agregam muita informação em ambientes externos, por apresentarem padrões que acontecem em outros lugares do ambiente, proporcionando assim um aumento no conjunto de informações a serem tratadas pelo modelo que não agregam muita informação relevante.

180° e assim pode-se perder informação ao utilizar apenas duas imagens, fazendo com que as amostras utilizadas não sejam completas. No entanto, a discussão sobre a influência da utilização de amostras completas no processo de caracterização de cenas será feita no Capítulo 5.

A escolha do FOV da câmera foi feita através da análise da possibilidade de ocorrer obstrução do campo de visão da mesma, uma vez que ambientes externos são extremamente dinâmicos. Por exemplo, considerando a cena apresentada na Figura 20, pode-se observar claramente na Figura 21 que a imagem capturada por uma câmera com maior FOV (à esquerda) apresenta informações importantes, presentes nos limites esquerdo e direito da imagem, que a imagem capturada por uma câmera com menor FOV (à direita) não apresenta, dado que um caminhão apareceu no campo de visão da mesma.

Esse tipo de situação ilustra bem o ganho de desempenho que pode ser obtido através da utilização de câmeras do tipo grande angular, pois mesmo que o modelo seja robusto o bastante para não ser prejudicado por um objeto não característico da cena, seria gasto tempo de processamento desnecessário para extrair características da imagem gerada pela câmera com menor FOV, pois estas seriam irrelevantes para o modelo.



Figura 20 – Exemplo de cena a ser caracterizada.



Figura 21 – Comparação entre imagens da mesma cena, porém capturadas por câmeras com FOVs diferentes. À esquerda pode-se observar a imagem obtida por uma câmera com aproximadamente 180° de FOV, e à direita é apresentada uma imagem que simula a diminuição do FOV da câmera.

Assim, uma vez gerado um conjunto de imagens através da captura de n amostras completas de um lugar do mapa, são extraídas as características SURF presentes neste conjunto³. Em seguida, são selecionadas apenas as características diferentes entre si para formar o conjunto característico deste lugar. Esta seleção é realizada pela eliminação de características que apresentem correspondências dentro do grupo de imagens do lugar.

Então, é efetuado o cálculo de frequência relativa das características presentes no conjunto de imagens, considerando como universo de ocorrência o conjunto selecionado anteriormente. Dessa forma, é gerado um histograma normalizado de características desse lugar que será utilizado como uma aproximação da PMF da variável L definida anteriormente. Logo, para caracterizar todo o ambiente em questão basta repetir esse procedimento para todos os lugares do mapa.

Para exemplificar esse processo, considere que um lugar definido por 3 amostras completas apresenta as características exibidas na Tabela 1, onde a presença do x na posição ix_j da tabela⁴ representa a ocorrência da característica i na amostra j . A função de frequência absoluta das características (FA), ou seja, o histograma das características desse lugar, é apresentada na Figura 22. Finalmente, para estimar a PMF desse lugar, basta normalizar a função de frequência absoluta das características, gerando assim a função de frequência relativa (FR) das k características desse lugar, como apresentado nas Equações 3.2 e 3.3 a seguir,

³ As quantidades de características extraídas dos lugares do mapa não são iguais, pois estas dependem das imagens de entrada que são submetidas ao algoritmo de extração SURF.

⁴ i representa o índice das linhas e j o índice das colunas.

Tabela 1 – Exemplo de ocorrência de características em 3 amostras de um lugar.

Característica	Amostra 1	Amostra 2	Amostra 3
1	x	x	x
2	x		x
3	x		x
4		x	
5		x	
6			x
7		x	x
8	x	x	x
9	x		
10	x		

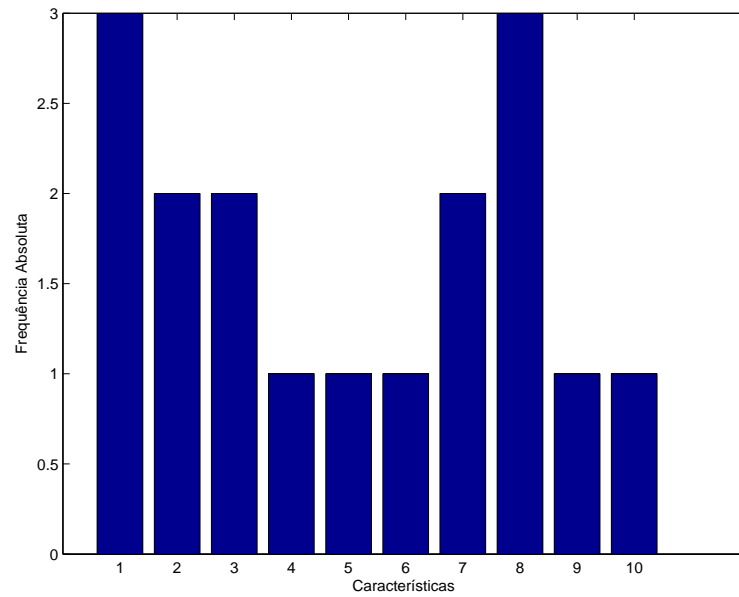


Figura 22 – Exemplo de função de frequência absoluta das características de um lugar.

$$PMF_L(L_i) = PMF_{s_i} \approx FR(s_i), \quad (3.2)$$

$$FR(s_i) = \frac{FA(s_i)}{\sum_{w=1}^k FA(s_{iw})}. \quad (3.3)$$

3.1.2 Determinação de Correspondências entre Lugares

Uma vez computadas as estimações das PMFs de todos os lugares do mapa, deseja-se calcular a probabilidade de uma amostra qualquer pertencer a esses lugares para determinar sua localização.

Por definição, para cada valor de entrada na PMF de uma variável aleatória é retornada a chance de que essa variável assuma o valor de entrada. No problema de localização, utilizou-se o valor de retorno da PMF_{s_i} de um lugar L_i como a probabilidade Pr de uma característica c_j qualquer ocorrer neste lugar, como definido a seguir

$$Pr_{s_i}(s_{iw} = c_j) = PMF_{s_i}(s_{iw} = c_j) \approx FR_{s_i}(s_{iw} = c_j). \quad (3.4)$$

Por exemplo, considere a PMF definida pela Tabela 1. Caso exista correspondência entre a característica c_j e a característica 1 de s_i , essa probabilidade é calculada pela Equação 3.5, como se segue

$$Pr_{s_i}(1) = PMF_{s_i}(1) \approx FR_{s_i}(1). \quad (3.5)$$

No entanto, é necessário calcular a probabilidade da amostra de um lugar qualquer pertencer aos lugares do mapa, para finalmente determinar a localização dessa amostra. Para tal, definiu-se a probabilidade de uma amostra c qualquer pertencer a um lugar L_i do mapa, pela soma dos valores da PMF_{s_i} desse lugar avaliada em características de s_i que apresentem correspondências com as características de c , como mostrado a seguir

$$Pr_{L_i}(c) = \sum_{c_j=1}^m PMF_{s_i}(s_{iw} = c_j) \approx \sum_{c_j=1}^m FR_{s_i}(s_{iw} = c_j), \quad (3.6)$$

onde c é uma amostra de um lugar qualquer, ou seja, conjunto de m características que representam este lugar; e c_j é uma característica do conjunto c .

Para justificar a Equação 3.6 pode-se analisar o espaço amostral da variável aleatória s_i nas seguintes situações:

- se uma amostra c contém todas as k características que representam um lugar L_i , ela apresentará probabilidade 1 de pertencer a este lugar;
- já se uma amostra c não contém nenhuma das k características que representam um lugar L_i ela apresentará probabilidade 0 de pertencer a este lugar;
- por fim, se uma amostra c contém uma quantidade parcial das k características que representam um lugar L_i , sua probabilidade de pertencer a este lugar será calculada segundo a Equação 3.6.

Por exemplo, considere o local L_i descrito na Tabela 1 e uma amostra (c) capturada em um lugar qualquer do mapa do ambiente ao qual pertence L_i . A Tabela 2 mostra a determinação de correspondências entre as características presentes na amostra c , e as características que determinam o domínio da PMF do lugar. Os números na segunda coluna da Tabela 2 representam as características da PMF do lugar, que são correspondentes às características da amostra c , presentes na primeira coluna. A Equação 3.7 representa a probabilidade da amostra c pertencer ao lugar L_i .

Tabela 2 – Correspondência das características da amostra c capturada, com as características s_i do lugar L_i descrito pela Tabela 1.

Característica da amostra c	Correspondência com s_i (domínio de FR_{s_i})
1	1
2	6
3	1
4	2
5	9

$$Pr_{L_i}(c) \approx FR_{s_i}(1) + FR_{s_i}(2) + FR_{s_i}(6) + FR_{s_i}(9) \quad (3.7)$$

Dado que é possível calcular a probabilidade de uma amostra qualquer c do ambiente pertencer a um lugar L_i do mapa, é necessário definir uma forma prática para determinar a localização da amostra no mapa. O método mais intuitivo é definir a localização da amostra $X(c)$ através do maior valor de probabilidade determinado entre todas as PMFs dos n lugares do mapa, segundo apresentado em

$$X(c) = \operatorname{argmax}_{L_i}(Pr_{L_1}(c), Pr_{L_2}(c), Pr_{L_3}(c), \dots, Pr_{L_n}(c)). \quad (3.8)$$

Contudo, esse método apresenta ordem de complexidade $\Theta(n)$ pois requer o cálculo das probabilidades da amostra c pertencer a todos os lugares do mapa. Utilizando uma abordagem bayesiana, não utilizada neste trabalho, que considere informações *a priori* é possível chegar a uma ordem de complexidade $O(n)$ para o método, uma vez que só seria necessário calcular as probabilidades para um subconjunto de n .

Logo, utilizou-se a definição de um método baseado em limiar. Analisando um conjunto de imagens, diferentes do conjunto utilizado para definir as PMFs do mapa, é definido um valor de referência vr para que seja possível dizer se a amostra c pertence ou não a um lugar do mapa L_i baseado no valor da probabilidade da amostra pertencer a

esse lugar $Pr_{L_i}(c)$. Se o valor de $Pr_{L_i}(c)$ for maior ou igual ao limiar vr , pode-se concluir que a amostra pertence ao lugar L_i , e caso contrário conclui-se que ela não pertence. Assim, o método de localização passa a apresentar ordem de complexidade $O(n)$, pois só é necessário calcular as probabilidades para todos os lugares do mapa no pior caso. A equação a seguir mostra o uso do limiar vr

$$X(c) = \begin{cases} \neg L_i & \text{se } Pr_{L_i}(c) < vr \\ L_i & \text{se } Pr_{L_i}(c) \geq vr. \end{cases} \quad (3.9)$$

Para realizar a determinação desse limiar utilizaram-se curvas ROC, pois além de apresentarem uma solução prática e rápida para avaliar o método, podem servir para determinação de limiares de forma precisa, como apresentado em [Sanchez-Gonzalez et al. \(2012\)](#). O critério utilizado para determinação do limiar foi a escolha de um valor que considere igualmente importantes as taxas de verdadeiro positivo (TPR) e falso positivo (FPR) que compõem as curvas ROC, ou seja, o ponto da curva que está mais próximo do canto superior direito do gráfico, como mostrado na Figura 23. Para avaliar o desempenho do método proposto são apresentados alguns testes no Capítulo 5.

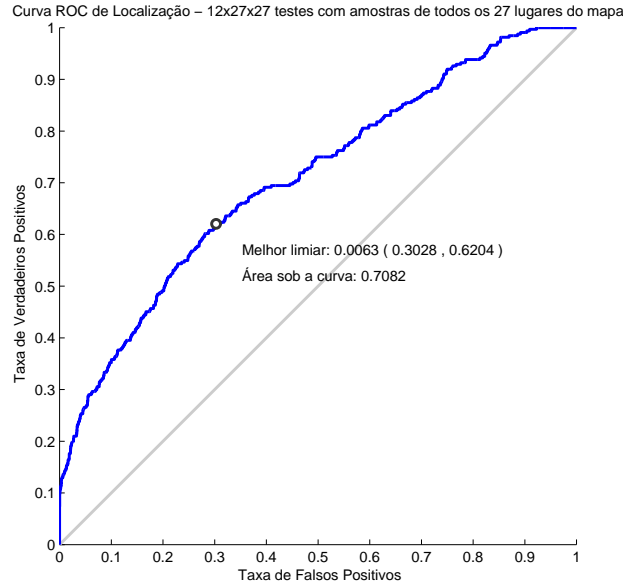


Figura 23 – Exemplo de determinação de limiar utilizado para testar a localização de uma amostra em lugar qualquer do mapa. O ponto circulado sobre a curva representa o limiar que gera melhores taxas de TPR e FPR.

3.2 Caracterização de Cenas Aplicada ao Problema de Reconhecimento de Objetos

O reconhecimento de objetos consiste em identificar objetos previamente caracterizados, ou seja, é um problema de classificação. Aplicando a mesma sequência de passos descrita na Seção 3.1 é possível realizar a caracterização de objetos, contudo é necessário definir uma amostra completa para o conjunto de objetos a ser caracterizado.

A determinação de uma amostra completa para o problema de reconhecimento de objetos depende das especificidades envolvidas no processo de reconhecimento dos objetos em questão. Para o objeto apresentado na Figura 19 são necessárias 8 imagens para compor uma amostra completa, considerando que o objeto sempre será apresentado para câmera de forma a mostrar apenas suas características externas. Por outro lado, para o controle remoto apresentado na Figura 24 são necessárias apenas 4 imagens para compor uma amostra completa.



Figura 24 – Exemplo de amostra completa para uma controle remoto.

O processo de geração das PMFs para caracterizar cada objeto do conjunto de objetos a ser reconhecido, segue a mesma sequência apresentada na Seção 3.1.1 e o método para realizar o reconhecimento é feito da mesma forma que o método proposto na Seção 3.1.2. A aplicação direta do modelo proposto para o reconhecimento de objetos se dá pelo fato de que não foram utilizadas informações métricas para caracterização dos lugares, ou seja, o modelo é todo baseado em características SURF que apresentam apenas informações de aparência das cenas. Para avaliar o desempenho do método proposto aplicado ao reconhecimento de objetos são apresentados alguns testes no Capítulo 5.

Parte IV

Experimentos

4 Metodologia de Avaliação

Nesse capítulo são apresentadas as métricas utilizadas para avaliar o modelo de caracterização de cenas proposto no Capítulo 3. Uma vez definido o modelo, é necessário avaliar o seu desempenho para definir se essa abordagem pode ser utilizada de forma robusta em métodos de localização visual autônoma, e se este também pode ser aplicado em métodos de reconhecimento de objetos. Para tal, foram utilizadas curvas ROC e matrizes de confusão para avaliar respectivamente, o modelo e o método de localização apresentados no Capítulo 3.

4.1 Curvas ROC

Utilizou-se a técnica gráfica de análise de desempenho de classificadores através de curvas ROC, pois segundo [Fawcett \(2006\)](#) esta técnica é útil para a avaliação do desempenho de classificadores, além de ser muito difundida na área de reconhecimento de padrões e aprendizagem de máquina.

Uma das justificativas para essa técnica de análise de desempenho ser bastante difundida, é o fato dela apresentar propriedades que a tornam invariante ao problema da desigualdade do número de amostras de teste entre as classes positivo e negativo (*class skew problem*). Ou seja, é possível mensurar a capacidade de distinção apresentada por um classificador, nesse caso um modelo de localização, sem depender da regra de classificação utilizada sobre este modelo. Logo, a utilização de curvas ROC aplicada ao problema de localização, serve para avaliar o modelo utilizado independentemente do método de localização utilizado sobre este.

Para isso, a construção de uma curva ROC é feita pelo cálculo das taxas de verdadeiros positivos e falsos positivos, geradas pela variação do limiar utilizado para classificar um conjunto de amostras de teste em duas classes (positivo e negativo). A definição de cada ponto da curva, é feita pelo cálculo das taxas de verdadeiros positivos, representada no eixo das ordenadas, e de falsos positivos, representada no eixo das abscissas, usando um limiar de classificação entre 0 e 1.

De forma prática, em uma abordagem conservadora que deseja minimizar as taxas de falsos positivos do classificador, um ponto P1 é melhor que outro P2, no espaço ROC, se P1 encontra-se mais próximo do ponto (0;1) (canto superior esquerdo) que P2 ([FAWCETT, 2006](#)), como pode ser visto na Figura 25, onde D é melhor que A, B, C e E. Assim, pode-se dizer que o ponto de uma curva ROC que está mais próximo do ponto (0;1) é o ponto de melhor desempenho do classificador. Então, este ponto é definido pelo melhor limiar de

classificação encontrado para o conjunto de amostras de teste utilizado, pois este limiar gera a maior taxa de verdadeiros positivos para a menor taxa de falsos positivos.

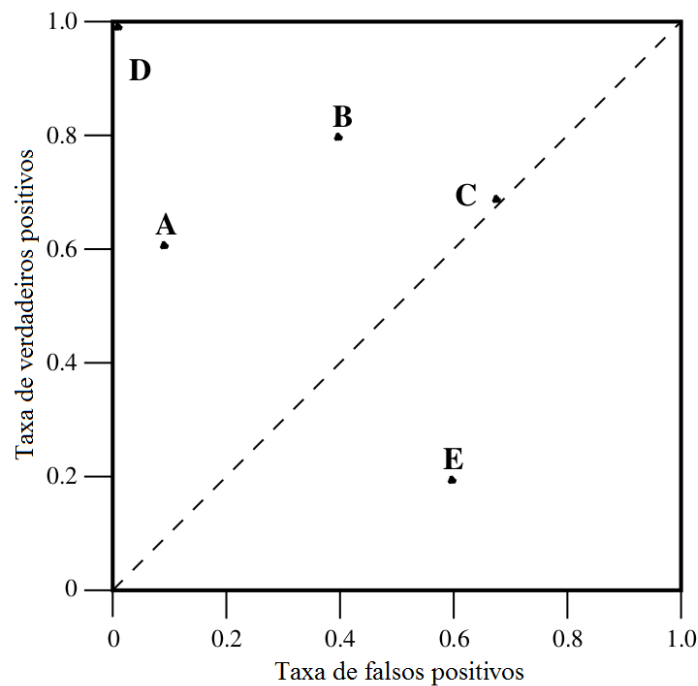


Figura 25 – Representação gráfica do espaço ROC. Quanto mais próximo do canto superior esquerdo do gráfico, melhor é o desempenho do classificador associado ao ponto. Fonte: [Fawcett \(2006\)](#).

Assim, pode-se comparar duas curvas ROC através dos seus pontos de melhor desempenho, como apresentado na Figura 26(a), onde a curva B representa claramente melhor desempenho do classificador se comparada à curva A. No entanto, quando se deseja comparar duas curvas que apresentem seus pontos de melhores desempenho idênticos ou semelhantes, não é possível compará-las de forma objetiva.

Então, pode-se definir qual dos classificadores é melhor através de um único escalar, utilizando a área sob a curva ROC, *Area Under Curve* (AUC) ([FAWCETT, 2006](#)). Quanto maior for a AUC de uma curva ROC, melhor será o desempenho do classificador. Dessa forma, mesmo na situação apresentada na Figura 26(b), na qual as curvas A e B apresentam seus pontos de melhor desempenho iguais, é possível afirmar que o classificador associado à curva B é melhor que o classificador associado à curva A, pois $AUC(B) > AUC(A)$ ([FAWCETT, 2006](#)).

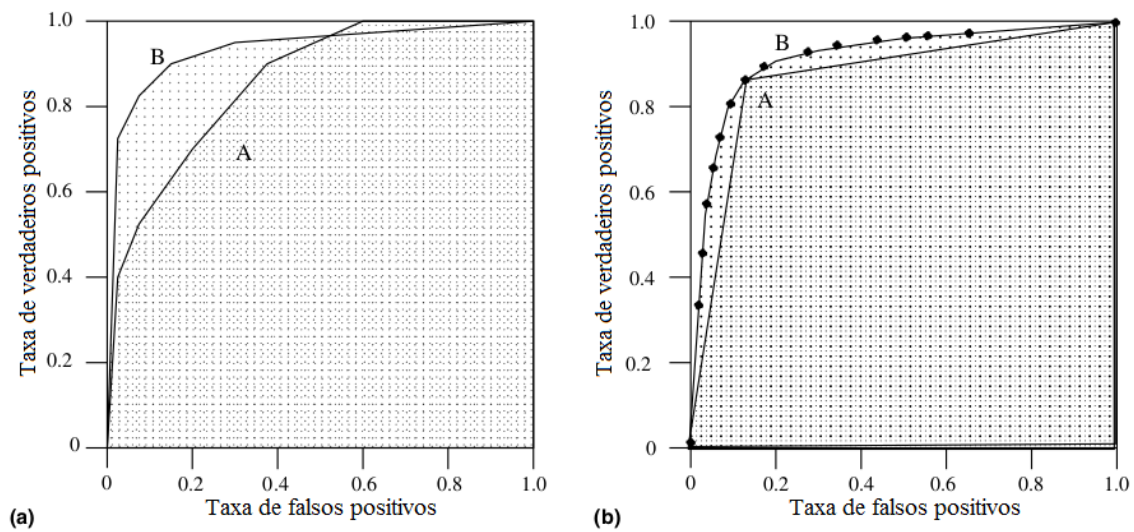


Figura 26 – Na Figura (a) é possível comparar as curvas sem utilizar a área sob elas. No entanto, em (b), a métrica AUC mostra-se necessária para que seja possível comparar as duas curvas. Fonte: [Fawcett \(2006\)](#).

Como a AUC de qualquer curva ROC é uma parte da área de um quadrado unitário, seu valor sempre estará entre 0 e 1. Entretanto, como um classificador aleatório produz no espaço ROC a linha diagonal entre os pontos (0;0) e (1;1), que apresenta $AUC = 0,5$, é desejável que um classificador obtenha $AUC > 0,5$ para que seja considerado realista ([FAWCETT, 2006](#)).

Todavia, não se pode considerar que todo classificador que apresente $AUC > 0,5$ é um bom classificador. Consequentemente, é necessário definir algumas faixas de AUC e suas respectivas classificações para que seja possível analisar o desempenho de um classificador de forma objetiva. Segundo [Sanchez-Gonzalez et al. \(2012\)](#) pode-se classificar o desempenho de classificadores através da AUC de sua curva ROC da seguinte forma:

- $AUC \leq 0,5$: ruim;
- $0,5 < AUC \leq 0,6$: pobre;
- $0,6 < AUC \leq 0,7$: razoável;
- $0,7 < AUC \leq 0,8$: aceitável;
- $0,8 < AUC \leq 0,9$: excelente;
- $0,9 < AUC \leq 1$: superior (*outstanding*);

Dessa forma, neste trabalho foram utilizados os critérios acima, para avaliar o desempenho do modelo apresentado no Capítulo 3. E o dado utilizado como entrada para

a construção da curva ROC de um lugar do mapa, é resultante da probabilidade de uma amostra qualquer pertencer a este lugar e da indicação se esta amostra realmente pertence ou não ao lugar em questão.

Por exemplo, se for calculada a probabilidade de 0,065 de uma amostra pertencer ao lugar L_x de um mapa, e esta amostra realmente pertencer a L_x o par (0,065;1) será utilizado como dado de entrada para a construção da curva ROC do lugar L_x . Isto significa que o classificador disse que existe 6,5% de chance desta amostra pertencer ao lugar L_x , e que ela realmente pertence. Caso esta amostra não pertencesse ao lugar L_x , seria utilizado o par (0,065;0) como dado de entrada da curva ROC.

Uma vez repetido esse processo para um conjunto de amostras, diferentes das amostras utilizadas para gerar a PMF do lugar, é gerado um conjunto de pares contendo a probabilidade do lugar e o resultado desejado da localização (0 ou 1); e então este conjunto é utilizado para gerar a curva ROC do lugar em questão.

Neste trabalho, para todos os testes efetuados, as curvas ROC foram geradas utilizando a seguinte metodologia:

- é capturado um conjunto de imagens de teste para cada lugar do mapa (com o mesmo número de amostras para cada lugar), garantindo um número mínimo de dias¹ entre essa captura e a captura do conjunto de imagens utilizado para definir a PMF do lugar. Além disso, todas as imagens de teste foram capturadas de posições, afastadas aproximadamente 4 metros, das posições de onde foram capturadas as imagens de caracterização. Pois, em ambientes dinâmicos, dificilmente vai ser possível voltar exatamente às posições métricas, das quais foram capturadas as imagens utilizadas para caracterização;
- são gerados os conjuntos de amostras de teste de cada lugar do mapa, a partir dos respectivos conjuntos de imagens de teste;
- para cada amostra de teste de um lugar, é calculada a probabilidade desta pertencer a todos os lugares do mapa, e então são gerados os pares correspondentes, formando assim o conjunto de pares do lugar;
- o passo anterior é repetido para os conjuntos de amostras de todos os lugares do mapa;
- para cada conjunto de pares de um lugar, é gerada uma curva ROC;
- os conjuntos de pares de todos os lugares do mapa são agrupados para formar o conjunto de pares do mapa, e a curva ROC correspondente é gerada.

¹ O número mínimo de dias variou conforme o teste, entretanto esses números são apresentados na Seção 5.2.

Gerando uma curva ROC por lugar (local), e também uma para o mapa como um todo (global), pode-se analisar o desempenho do modelo, respectivamente, de forma local e global. Consequentemente, foi possível analisar as especificidades de cada lugar e a sua influência no desempenho do modelo.

4.2 Matriz de Confusão

Utilizando a análise de desempenho através de curvas ROC, é possível realizar uma análise objetiva através do ponto de melhor desempenho e da AUC das curvas. Contudo, não é possível visualizar o padrão dos erros entre as classes do problema, neste caso os lugares do mapa. Logo, utilizou-se neste trabalho a análise por matriz de confusão para discutir os resultados tanto do modelo quanto do método de localização.

Matrizes de confusão servem para identificar os erros gerados por classificadores, e também para definir quais classes estão sendo confundidas. Sendo assim, para montar a matriz de confusão, é necessário submeter um conjunto de amostras de teste, previamente classificadas de forma correta, à avaliação do classificador. Então, a matriz de confusão apresentará em suas linhas, as classes às quais realmente pertencem as amostras de teste; e nas colunas, a indicação (absoluta ou percentual) do classificador para as amostras de teste de cada classe.

Por exemplo, considere um sistema que classifica imagens nos seguintes grupos: cachorro, gato, peixe e pessoa. Neste caso, é apresentado na Tabela 3 a matriz de confusão do classificador, gerada para um conjunto de 40 imagens, sendo 10 pertencentes a cada grupo.

Tabela 3 – Exemplo de matriz de confusão.

	Cachorro	Gato	Peixe	Pessoa
Cachorro	6	4	0	0
Gato	3	7	0	0
Peixe	0	0	10	0
Pessoa	0	0	0	10

Analisando a Tabela 3 é possível perceber que o suposto classificador identifica de forma precisa imagens de pessoas e de peixes, pois todas as 10 imagens de peixes foram classificadas como peixes e todas as 10 imagens de pessoas foram classificadas como pessoas. No entanto, existe um certo grau de confusão do classificador em relação às classes cachorro e gato, pois das 10 imagens de cachorros 6 foram classificadas como cachorro e 4 como gato, e das 10 imagens de gatos 7 foram classificadas como gato e 3 como cachorro.

Para um classificador ser considerado perfeito, uma característica importante da sua matriz de confusão, que tenha sido gerada para um conjunto de amostras de teste com N amostras por classe, é que esta deve ser diagonal com todos os valores da diagonal iguais a N . Pois dessa forma, nenhuma das amostras será classificada de forma errada.

Logo, neste trabalho foram utilizadas matrizes de confusão para analisar tanto o desempenho do modelo, quanto do método de localização apresentados no Capítulo 3. Todavia, é importante ressaltar que as matrizes utilizadas para avaliação de desempenho foram denominadas de matrizes de confusão*, pois a soma de seus valores presentes em cada linha não é igual à quantidade de amostras de teste por classe.

Isso acontece, porque estas matrizes foram geradas para a regra de classificação local apresentada na Equação 3.9, que define o seguinte: se o valor da probabilidade da amostra de teste for superior ao limiar do lugar em hipótese, considera-se que a amostra pertence a este lugar, caso contrário ela não pertence. Dessa forma, é possível que uma amostra seja considerada pertencente a mais de um lugar do mapa, uma vez que essa regra de decisão é local e não global como a maioria dos classificadores.

5 Avaliação do Modelo

Neste Capítulo são apresentados alguns testes para avaliar o modelo em um ambiente externo e dinâmico. Inicialmente, são apresentadas as análises efetuadas para determinação da quantidade de imagens necessárias para caracterizar os lugares de um ambiente externo, que apresentam pouca similaridade entre si, de forma a manter uma boa relação entre custo computacional e a qualidade da caracterização.

Em seguida são apresentados os principais resultados deste trabalho, obtidos através de experimentos elaborados em um ambiente composto por 28 lugares que apresentam uma grande quantidade de características em comum. Além de avaliar o desempenho do modelo proposto são analisados os resultados da aplicação do método de localização apresentado no Capítulo 3, que utiliza limiares definidos pelas curvas ROC de localização dos lugares do ambiente.

Com a utilização desse ambiente foi possível analisar a influência da quantidade de características extraídas dos lugares, da iluminação, da utilização de amostras incompletas, e do descarte de características das PMFs no processo de caracterização dos lugares. Além disso, foram apresentadas discussões sobre o comportamento do modelo em regiões de transição entre os lugares do mapa, e sobre as características temporais dos algoritmos desenvolvidos.

Por fim, são apresentados resultados da aplicação do modelo de caracterização de cenas ao reconhecimento de objetos, com o objetivo de mostrar a possibilidade de utilização do modelo em aplicações diferentes da localização em ambientes externos. Para isso, foi utilizado um conjunto formado por 5 objetos diferentes, caracterizados sobre um fundo branco e reconhecidos sobre um fundo diferente deste.

5.1 Número de Imagens para Caracterização

O primeiro parâmetro a ser determinado, para que seja possível utilizar o modelo na localização em ambientes externos, é a quantidade de imagens a ser utilizada para caracterizar cada lugar do mapa, pois a resolução utilizada em todos os testes foi fixada na máxima resolução da câmera, 1280x480¹.

Os estudos inicialmente realizados sobre este assunto (OLIVEIRA; VASSALLO; SALLES, 2011) mostraram que, mesmo em um mapa com 3 lugares distintos e que não apresentam visada direta² entre si, o aumento do número de amostras utilizadas para

¹ Utilizou-se a máxima resolução, pois em testes preliminares observou-se uma diminuição do índice AUC da curva ROC do mapa, gradativa à diminuição da resolução das imagens.

² A presença de lugares que apresentam visada direta entre si, torna o processo de localização mais

caracterizar os lugares, gera uma melhora no desempenho atingido pelo modelo. Entretanto, essa melhora apresenta um limite, a partir do qual o desempenho diminui com o aumento da quantidade de amostras, pois além de aumentar a ocorrência de características que representem bem o lugar, aumenta-se a ocorrência de características comuns a todos os lugares do mapa, que geram falsas correspondências, como padrões de céu e vegetação.

Assim, foi contruído um mapa com 5 lugares distintos, caracterizados por amostras completas, e que não apresentam visada direta entre si, como pode ser visto na Figura 27, a fim de determinar uma quantidade de amostras que seja suficiente para caracterizar os lugares de forma minimamente robusta.

Para garantir essa robustez, as imagens utilizadas para montar a PMF de cada lugar e suas respectivas imagens de teste, foram capturadas com um intervalo de tempo de 3 dias, e em todos os lugares foi garantida uma distância mínima de aproximadamente 4 metros, entre o ponto de captura de uma imagem de caracterização e o ponto de captura de sua respectiva imagem de teste.

Estas restrições foram feitas, pois em testes preliminares observou-se que para amostras de teste capturadas sob condições de iluminação parecidas e no mesmo ponto, o desempenho mínimo do modelo foi excelente ($AUC > 0,8$).



(a) Lugar L0.



(b) Lugar L1.



(c) Lugar L2.



(d) Lugar L3.



(e) Lugar L4.

Figura 27 – Imagens de exemplo dos lugares do mapa.

difícil, pois dependendo da distância que os separa, podem ser geradas falsas correspondências entre as características, o que pode afetar o desempenho do modelo se a caracterização não for efetuada de forma correta. Este assunto será abordado na Seção 5.5.

Após realizar o processo descrito na Seção 4.1, foram geradas curvas ROC para diferentes quantidades de amostras de caracterização dos lugares, sempre utilizando 13 amostras incompletas³ para teste, compostas por apenas uma imagem. Após análise da Tabela 4, observa-se que os resultados apresentados em Oliveira, Vassallo e Salles (2011) foram confirmados, pois o aumento do número de amostras de caracterização melhora o desempenho de classificação até o limite de 3 amostras, neste caso.

Além disso, foi possível obter resultado pelo menos excelente, tanto local quanto global, para todas as quantidades de amostras completas de caracterização testadas, segundo a classificação apresentada na Seção 4.1. As curvas ROC do melhor resultado podem ser vistas na Figura 28.

Tabela 4 – Relação entre o índice AUC e o número de amostras completas de caracterização.

Número de Amostras de Caracterização	Índice AUC
1	0.8891
2	0.8939
3	0.9015
4	0.8972

³ Não utilizaram-se amostras de teste completas pois em testes preliminares foi possível comprovar que para pequenos intervalos de tempo entre captura de amostras de caracterização e teste, a utilização de amostras de teste completas gera aproximadamente 100% de AUC tanto para curvas ROC locais quanto globais.

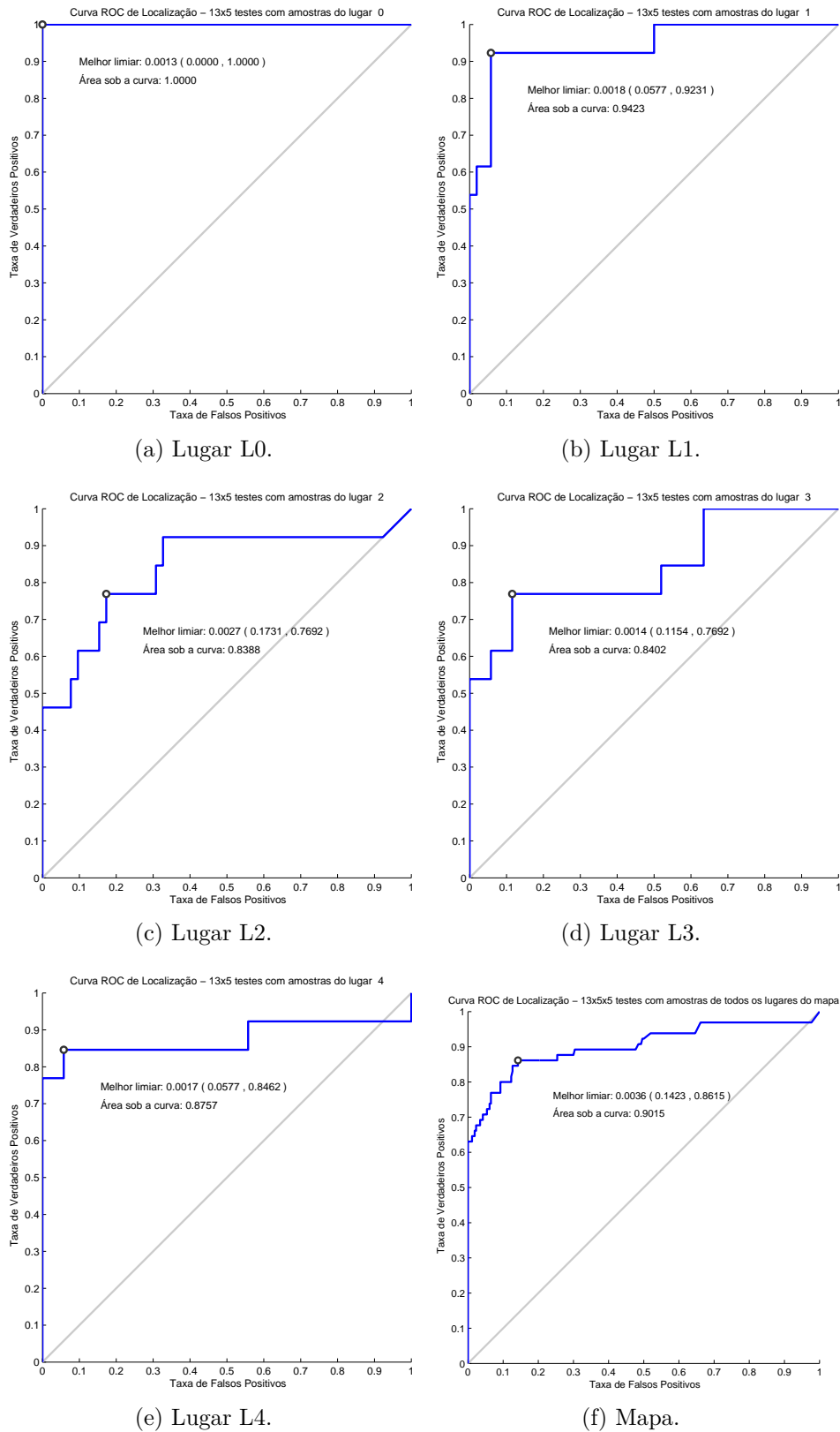


Figura 28 – Curvas ROC dos lugares e do mapa, geradas por 3 amostras completas de caracterização e 13 amostras simples (1 imagem apenas) de teste por lugar, totalizando 65 testes por lugar e 325 para o mapa.

É possível perceber que não existe uma uniformidade no desempenho do modelo, pois a forma da curva, o seu ponto de melhor desempenho e consequentemente o índice AUC é diferente para cada lugar. Uma explicação para tal, é feita pela observação de que cada lugar apresenta características diferentes e em quantidades diferentes, que podem ser mais fortes com relação a sua singularidade ou mais fracas, caracterizando assim a inserção de ruído nas PMFs dos lugares.

Como pode-se observar na Tabela 5, o lugar com melhor desempenho, L0, apresenta o menor número de características. Esse fato pode ser justificado pois uma PMF que apresente poucas características em relação às PMFs dos demais lugares do mapa, apresentará grandes probabilidades de ocorrência de características fortes. Isso acontece, pois a quantidade total de características utilizada para normalizar o histograma do lugar será pequena se comparada aos outros lugares do mapa.

Tabela 5 – Quantidade de características extraídas dos lugares do mapa.

Identificação	Número de Características
L0	3759
L1	13434
L2	10656
L3	6943
L4	9215
Mapa	44007

Logo, amostras de um lugar do mapa que possui poucas características comparado aos demais, terão altas probabilidades de pertencer a este lugar, sendo localizadas de forma correta. Por outro lado, amostras dos demais lugares podem ter altas probabilidades de pertencer a este mesmo lugar, por mais que apresentem poucas correspondências erradas, sendo assim localizadas de forma incorreta. Para comprovar esse fato, pode-se analisar a matriz de confusão* do mapa apresentada na Tabela 6.

Para o lugar que apresenta menor quantidade de características do mapa, L0, tem-se 34% dos testes indicando L0, para amostras de teste pertencentes a outros lugares do mapa, linhas 3, 4, 5 e 6 da segunda coluna; e 100% dos testes indicando L0, para amostras pertencentes a L0, linha 2 da segunda coluna. Já para o lugar com a maior quantidade de características do mapa, L1, tem-se 25% dos testes indicando L1, para amostras de teste pertencentes a outros lugares do mapa, linhas 2, 4, 5 e 6 da terceira coluna; e 100% dos testes indicando L1, para amostras pertencentes a L1, linha 3 da terceira coluna.

Tabela 6 – Matriz de confusão* do mapa.

	L0	L1	L2	L3	L4
L0	13	1	10	11	10
L1	11	13	3	8	8
L2	12	0	12	6	4
L3	11	4	7	11	11
L4	13	8	5	8	12

Apesar do excelente resultado, não se pode garantir que esse desempenho do modelo se mantenha quando o mapa apresentar um número maior de lugares, e/ou apresentar lugares com visada direta entre si. Por outro lado, também não é possível afirmar que 3 amostras por lugar é o mínimo necessário para alcançar um bom desempenho do modelo. Por isso, faz-se necessária a execução de experimentos mais extremos, que serão apresentados na Seção 5.2.

5.2 O Problema de Localização

Na Seção 5.1 pôde-se concluir que, para um mapa com poucos lugares que não apresentam visada direta entre si, o modelo obteve um bom desempenho através da utilização de 3 amostras completas na caracterização dos lugares. Sendo assim, para que fosse possível confirmar esse desempenho, foi executado um experimento sem as restrições impostas no experimento da Seção 5.1.

Foi construído um mapa com 28 lugares⁴ distintos, caracterizados por amostras completas, dos quais alguns apresentavam visada direta entre si (L7, L25, L26, L22, L23, L6, L21, L20, L5, L17, L12, L4, L13, L11, L19, L14, L15, L16, L27, L2), como pode ser visto na Figura 29.

⁴ Apesar de terem sido capturadas imagens para caracterização de 28 lugares, o lugar L18 não foi considerado neste experimento por ter sido isolado devido a obras, e consequentemente não ter sido possível capturar imagens para compor amostras de teste completas.

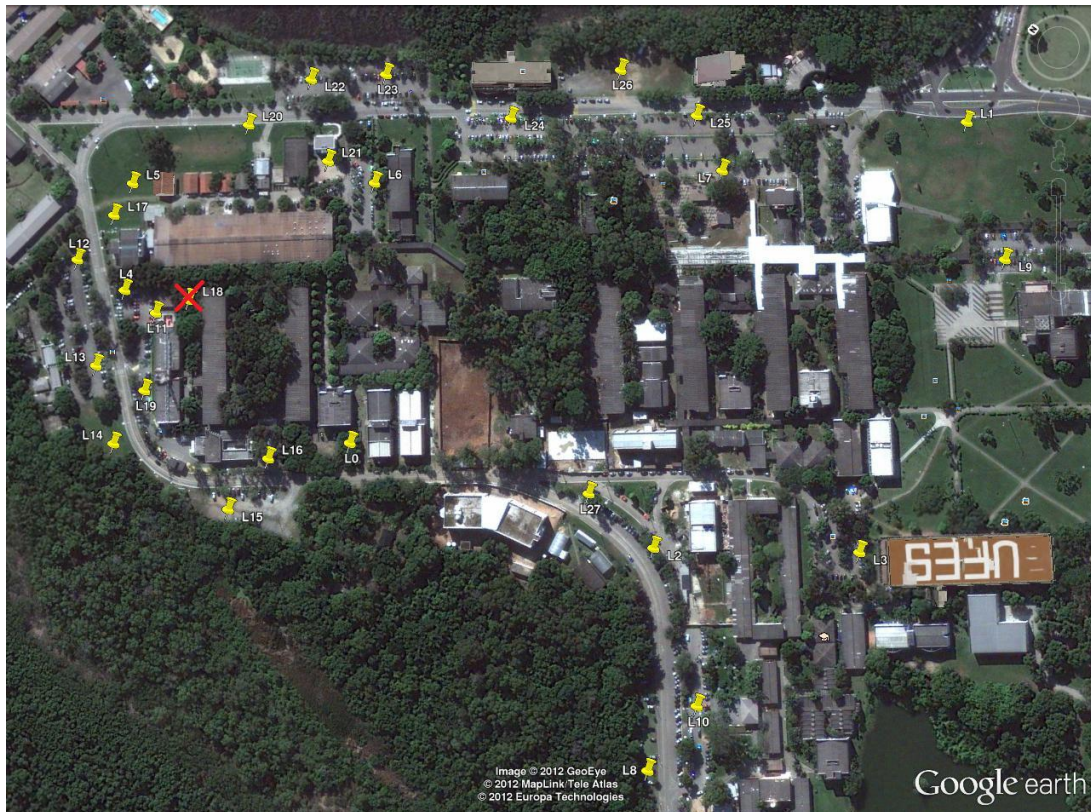


Figura 29 – Mapa do ambiente UFES, do qual foram utilizados, neste experimento, 27 dos seus 28 lugares representados pelos marcadores amarelos.

Para que fosse possível garantir a robustez do sistema, utilizou-se um intervalo de tempo bem maior, se comparado ao experimento da Seção 5.1, entre a captura das imagens utilizadas para compor as amostras de caracterização, e a captura das imagens utilizadas para compor as amostras de teste; e em todos os lugares foi garantida uma distância mínima de aproximadamente 4 metros, entre o ponto de captura de uma imagem de caracterização e o ponto de captura de sua respectiva imagem de teste.

Ao analisar a Tabela 7, é possível perceber que o valor mínimo desse intervalo foi de pelo menos 1 mês, e em períodos distintos do dia. Todavia, em mais da metade dos lugares esse intervalo foi de pelo menos 5 meses, mostrando assim que as imagens utilizadas no experimento foram capturadas com o objetivo de aproveitar as extremas mudanças de aparência que podem ocorrer nos lugares, uma vez que estes são dinâmicos e possuem grande variação na presença de objetos.

Para mostrar a aparência dos lugares do mapa, a variação de iluminação capturada e a variabilidade de objetos nos mesmo, são apresentadas na Figura 30, imagens exemplo das amostras utilizadas para caracterizar os lugares, e na Figura 31, imagens exemplo das amostras utilizadas para testar o modelo.

Tabela 7 – Períodos de captura das amostras de teste e de caracterização, utilizadas no experimento. Na segunda e terceira colunas temos o dia, a hora e os minutos de início e fim da captura, e na quarta coluna temos o intervalo de tempo mínimo (em meses) entre a captura das imagens de caracterização e das imagens de teste.

	Captura das Amostras de Caracterização	Captura das Amostras de Teste	Intervalo Mínimo (meses)
L0	17.01.2012 / 11:00 - 11:05	25.06.2012 / 12:51 - 12:56	5
L1	28.01.2012 / 17:22 - 17:25	13.03.2012 / 12:06 - 12:09	1
L2	17.01.2012 / 11:10 - 11:14	19.07.2012 / 15:47 - 15:52	6
L3	17.01.2012 / 11:38 - 11:40	19.07.2012 / 16:03 - 16:06	6
L4	12.01.2012 / 17:09 - 17:12	13.03.2012 / 11:44 - 11:48	2
L5	28.01.2012 / 16:53 - 16:56	13.03.2012 / 13:54 - 13:59	1
L6	28.01.2012 / 17:13 - 17:16	22.03.2012 / 14:47 - 14:50	1
L7	12.01.2012 / 11:52 - 11:57	19.07.2012 / 15:36 - 15:38	6
L8	17.01.2012 / 11:21 - 11:24	06.08.2012 / 16:52 - 17:00	6
L9	12.01.2012 / 17:48 - 17:50	28.02.2012 / 14:50 - 14:55	1
L10	17.01.2012 / 11:28 - 11:32	05.08.2012 / 16:41 - 16:43	6
L11	12.01.2012 / 16:58 - 17:02	22.03.2012 / 14:55 - 15:00	2
L12	12.01.2012 / 17:18 - 17:23	25.06.2012 / 13:04 - 13:09	5
L13	12.01.2012 / 17:28 - 17:30	26.06.2012 / 12:52 - 12:56	5
L14	12.01.2012 / 17:35 - 17:39	26.06.2012 / 13:00 - 13:05	5
L15	16.01.2012 / 18:04 - 18:10	26.06.2012 / 13:22 - 13:25	5
L16	28.01.2012 / 16:35 - 16:39	26.06.2012 / 12:40 - 12:44	4
L17	28.01.2012 / 16:46 - 16:49	26.06.2012 / 13:11 - 13:15	4
L18	14.03.2011 - 22.03.2011 / 12:00 - 17:00	-	-
L19	16.01.2012 / 17:55 - 17:59	27.07.2012 / 11:33 - 11:39	6
L20	28.01.2012 / 16:59 - 17:02	27.07.2012 / 11:44 - 11:49	5
L21	28.01.2012 / 17:07 - 17:10	27.07.2012 / 11:53 - 11:57	5
L22	12.01.2012 / 12:39 - 12:42	09.04.2012 / 16:30 - 16:34	2
L23	12.01.2012 / 12:32 - 12:35	09.04.2012 / 16:40 - 16:44	2
L24	12.01.2012 / 12:21 - 12:24	09.04.2012 / 16:50 - 16:54	2
L25	12.01.2012 / 12:06 - 12:08	05.08.2012 / 16:25 - 16:30	6
L26	12.01.2012 / 12:14 - 12:17	09.04.2012 / 17:02 - 17:11	2
L27	16.01.2012 / 18:17 - 18:19	27.07.2012 / 11:07 - 11:11	6



(a) Lugar L0.



(b) Lugar L1.



(c) Lugar L2.



(d) Lugar L3.



(e) Lugar L4.



(f) Lugar L5.



(g) Lugar L6.



(h) Lugar L7.



(i) Lugar L8.



(j) Lugar L9.



(k) Lugar L10.



(l) Lugar L11.



(m) Lugar L12.



(n) Lugar L13.



(o) Lugar L14.



(p) Lugar L15.



(q) Lugar L16.



(r) Lugar L17.



(s) Lugar L18.



(t) Lugar L19.



(u) Lugar L20.



(v) Lugar L21.



(w) Lugar L22.



(x) Lugar L23.



(y) Lugar L24.



(z) Lugar L25.



(aa) Lugar L26.



(ab) Lugar L27.

Figura 30 – Imagens exemplo das amostras de caracterização dos lugares do mapa.



(a) Lugar L0.



(b) Lugar L1.



(c) Lugar L2.



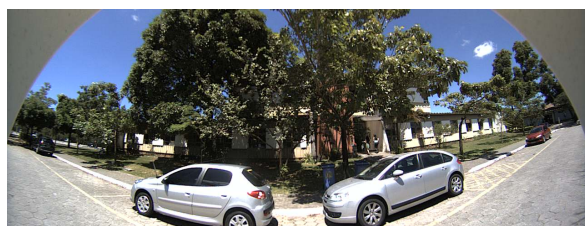
(d) Lugar L3.



(e) Lugar L4.



(f) Lugar L5.



(g) Lugar L6.



(h) Lugar L7.



(i) Lugar L8.



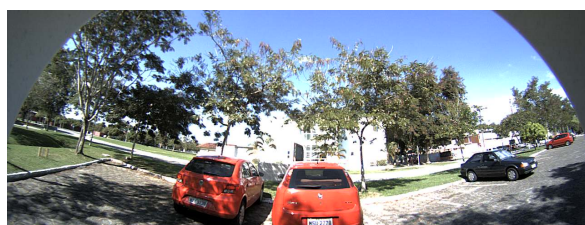
(j) Lugar L9.



(k) Lugar L10.



(l) Lugar L11.



(m) Lugar L12.



(n) Lugar L13.



(o) Lugar L14.



(p) Lugar L15.



(q) Lugar L16.



(r) Lugar L17.



(s) Lugar L19.



(t) Lugar L20.



(u) Lugar L21.



(v) Lugar L22.



(w) Lugar L23.



(x) Lugar L24.



(y) Lugar L25.



(z) Lugar L26.



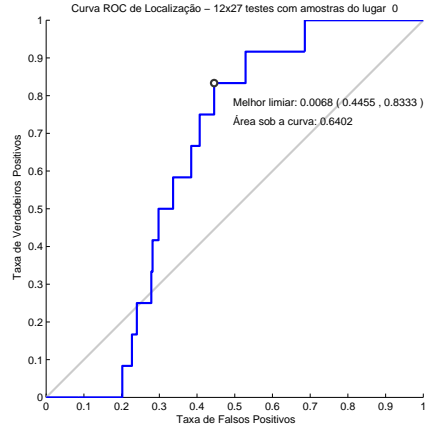
(aa) Lugar L27.

Figura 31 – Imagens exemplo das amostras de teste dos lugares do mapa.

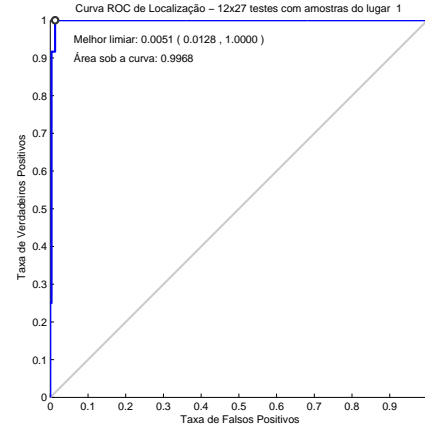
5.2.1 Análise Geral

Após realizar o processo descrito na Seção 4.1, foram geradas curvas ROC para diferentes quantidades de amostras de caracterização dos lugares, partindo de 3 como foi definido no experimento apresentado na Seção 5.1, e sempre utilizando 12 amostras completas para teste. Somente a partir da utilização de 10 amostras completas para

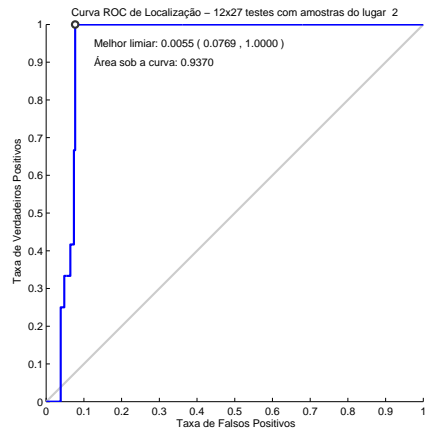
caracterização, foi possível obter um resultado aceitável de forma global, como pode ser visto na Figura 32.



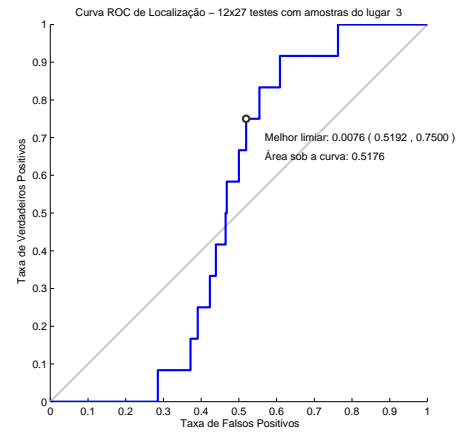
(a) Lugar L0.



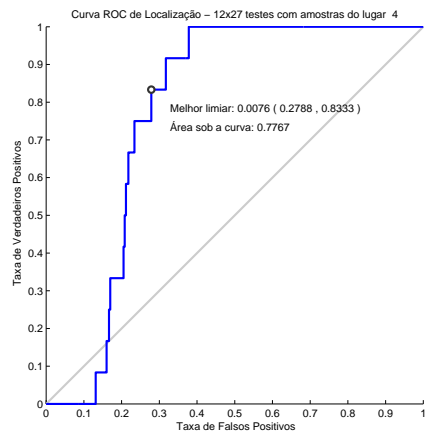
(b) Lugar L1.



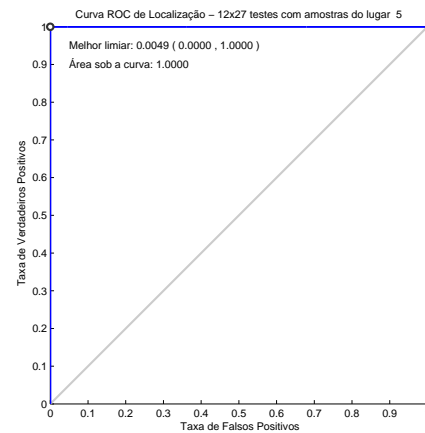
(c) Lugar L2.



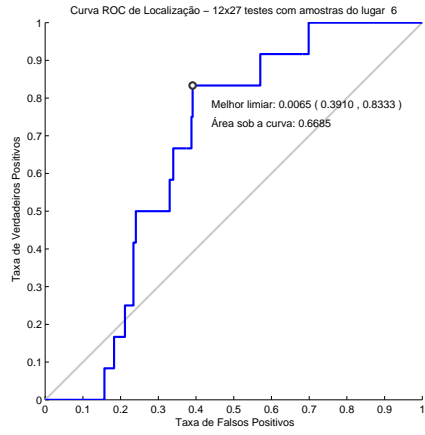
(d) Lugar L3.



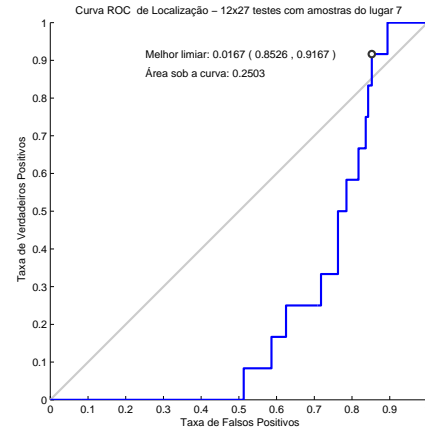
(e) Lugar L4.



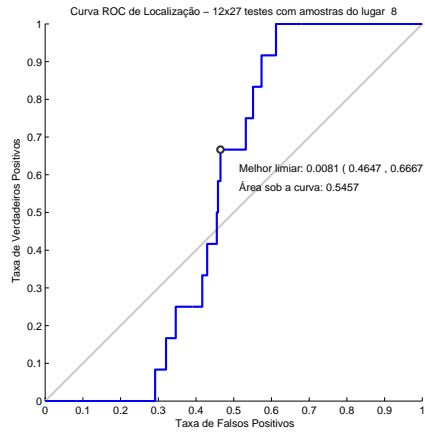
(f) Lugar L5.



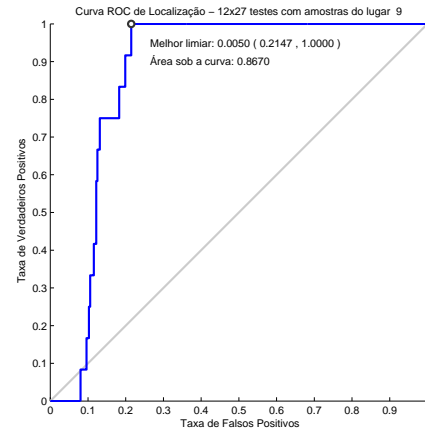
(g) Lugar L6.



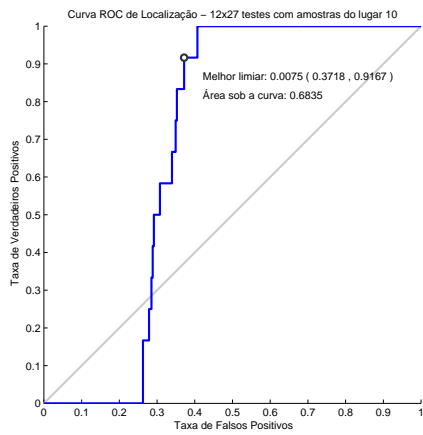
(h) Lugar L7.



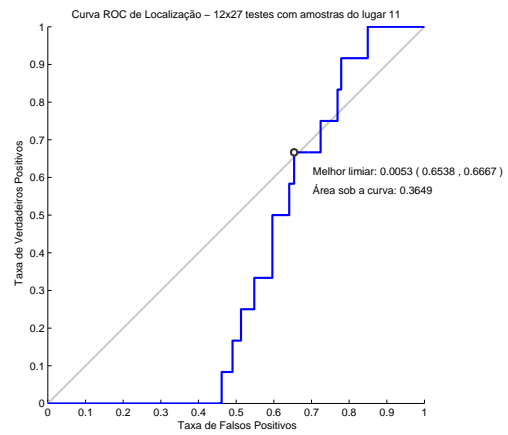
(i) Lugar L8.



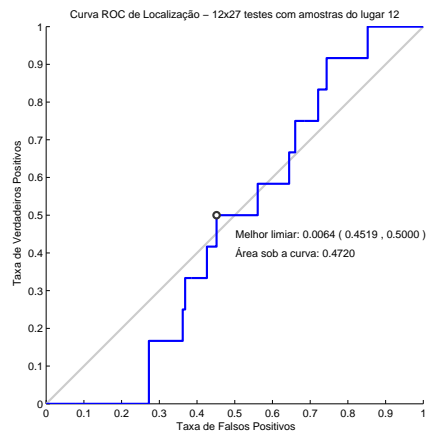
(j) Lugar L9.



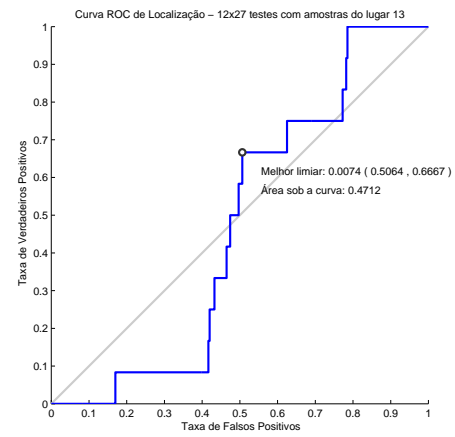
(k) Lugar L10.



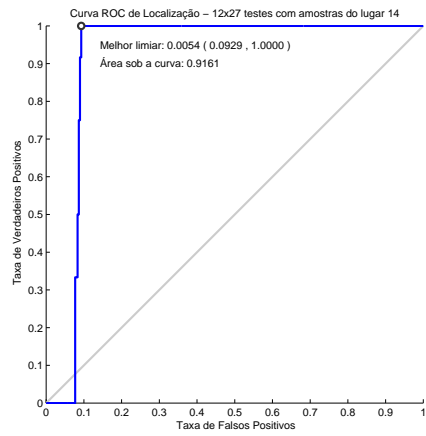
(l) Lugar L11.



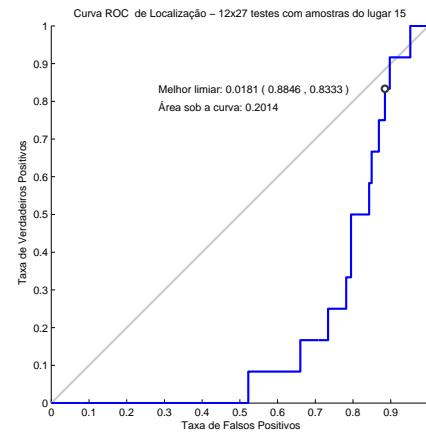
(m) Lugar L12.



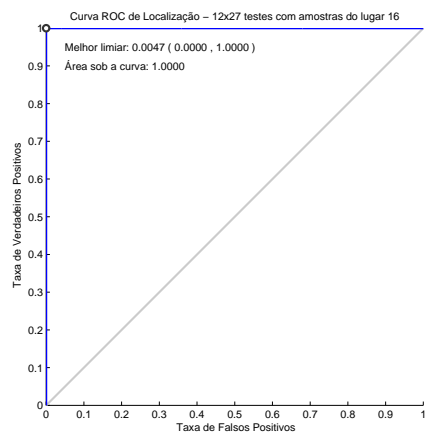
(n) Lugar L13.



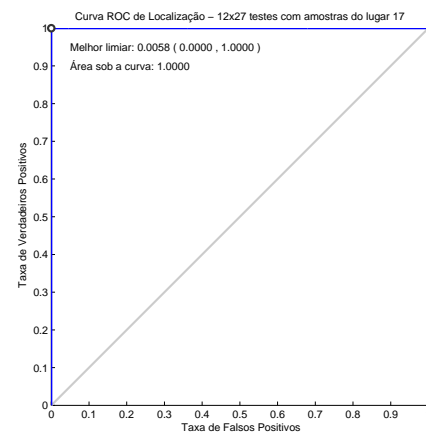
(o) Lugar L14.



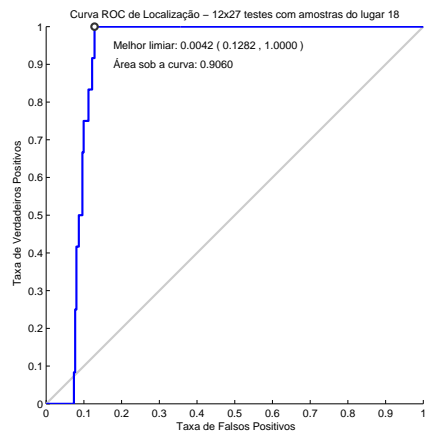
(p) Lugar L15.



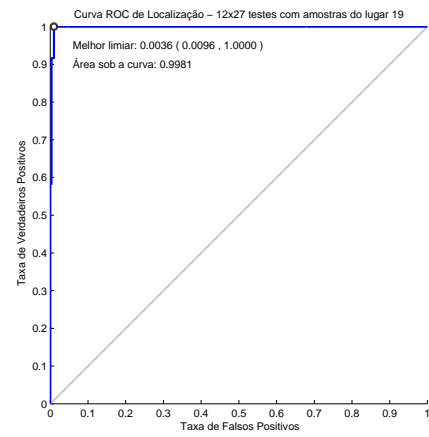
(q) Lugar L16.



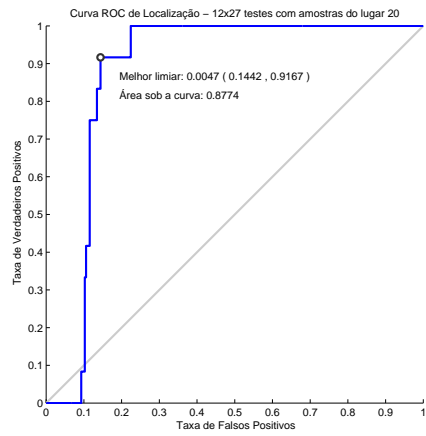
(r) Lugar L17.



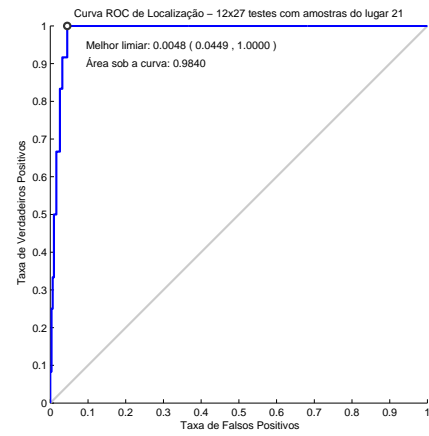
(s) Lugar L19.



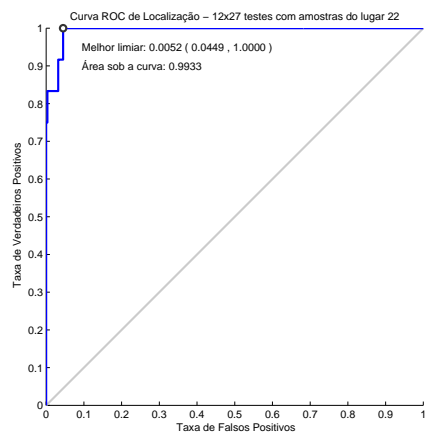
(t) Lugar L20.



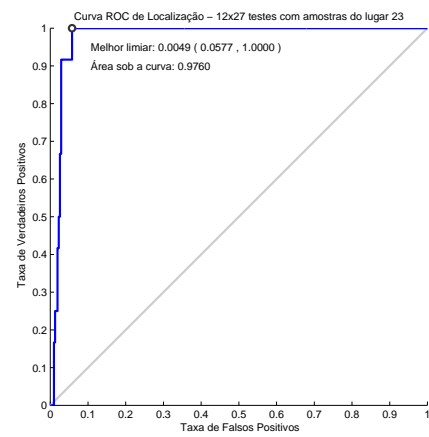
(u) Lugar L21.



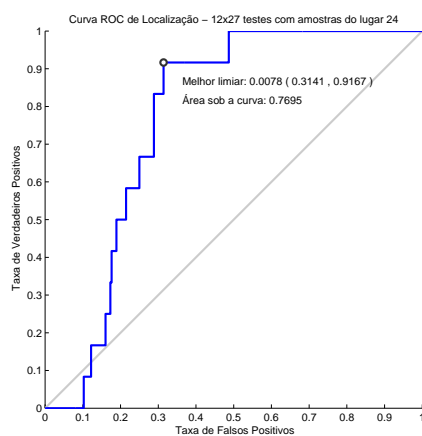
(v) Lugar L22.



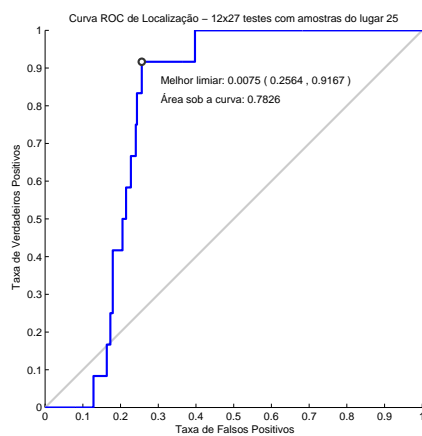
(w) Lugar L23.



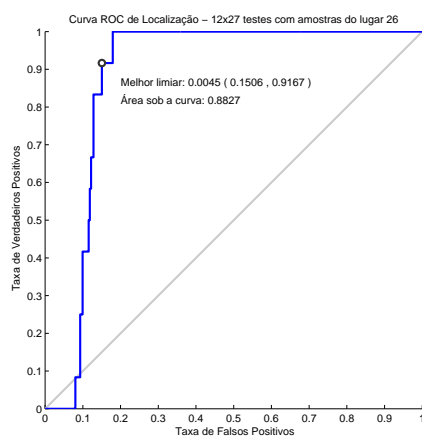
(x) Lugar L24.



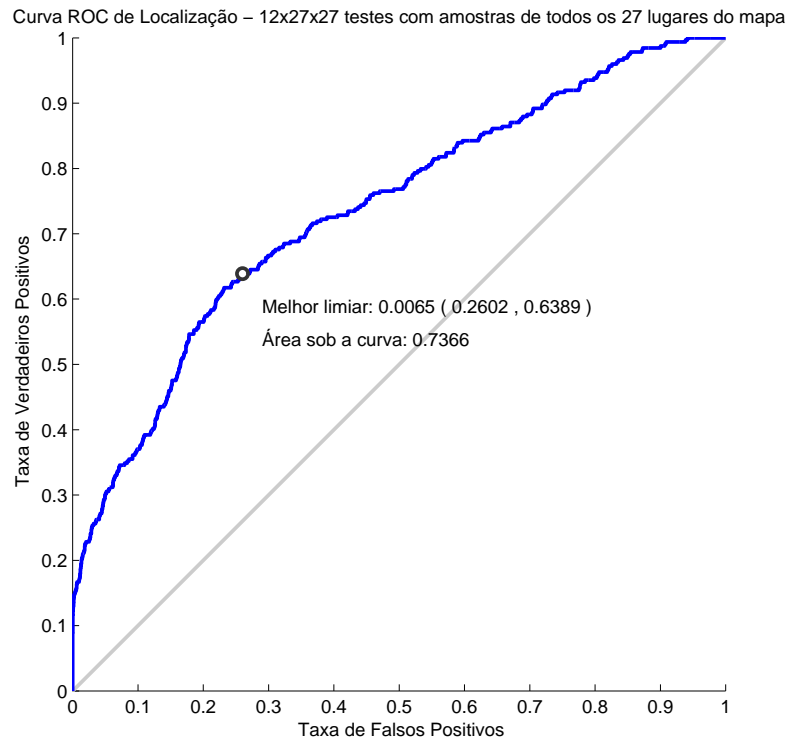
(y) Lugar L25.



(z) Lugar L26.



(aa) Lugar L27.



(ab) Mapa.

Figura 32 – Curvas ROC dos lugares e do mapa, geradas por 10 amostras completas de caracterização e 12 amostras completas de teste por lugar, totalizando 324 testes por lugar e 8748 para o mapa.

Percebe-se que o desempenho não uniforme entre os lugares do mapa foi mantido, se comparado ao experimento apresentado na Seção 5.1. Contudo, o excelente desempenho global não se confirmou pois o índice AUC classifica o modelo como aceitável segundo os critérios apresentados na Seção 4.1. Porém, em 52% dos lugares o modelo apresentou resultado excelente ou superior (L1, L2, L5, L9, L14, L16, L17, L19, L20, L21, L22, L23, L24 e L27), o que mostra a possível existência de condições sob as quais o modelo apresente desempenho global superior ou excelente.

5.2.2 Influência da Quantidade de Características

Com objetivo de analisar a existência dessas condições foi gerada a Tabela 8, que exhibe as quantidades de características extraídas de cada lugar do mapa, e a matriz de confusão* do experimento, apresentada na Tabela 9.

Tabela 8 – Quantidade de características extraídas dos lugares do mapa.

Identificação	Número de Características
L0	91551
L1	27259
L2	52943
L3	88423
L4	53223
L5	21162
L6	45741
L7	71428
L8	60079
L9	67129
L10	60915
L11	72389
L12	59812
L13	52816
L14	48323
L15	58312
L16	57573
L17	32160
L19	68607
L20	19574
L21	49141
L22	43662
L23	39076
L24	44189
L25	46937
L26	30764
L27	45055
Mapa	1408243

O número médio de características extraídas dos lugares que apresentaram desempenho abaixo do aceitável foi de 66146, já nos demais lugares, onde o desempenho mínimo foi aceitável, a média ficou em 43928. Ou seja, nos lugares que apresentaram desempenho abaixo do aceitável, a quantidade de características extraídas é 33,6% maior que nos demais, o que confirma o resultado apresentado na Seção 5.1.

Assim, lugares que apresentem uma quantidade significativamente maior de características em relação aos demais lugares do mapa, tendem a apresentar menor desempenho de classificação ao utilizar o modelo, e vice-versa. Novamente é possível confirmar essa

afirmativa através da matriz de confusão* do mapa, apresentada na Tabela 9.

Dos 52% dos lugares que apresentaram resultado excelente ou superior, os quais também apresentaram menor número de características, apenas 21% (L14, L21 e L27) não indicaram 100% dos testes corretamente, para amostras de testes pertencentes aos respectivos lugares. Para estes 3 lugares tal indicação foi correta para 92% dos teste, ou seja, 11 de 12.

Analisando o lugar com menor número de características, L20, tem-se 84% dos testes indicando L20, para amostras de teste pertencentes a outros lugares do mapa; e 100% dos testes indicando L20, para amostras pertencentes a L20. Já para o lugar com a maior quantidade de características do mapa, L0, tem-se 15% dos testes indicando L0, para amostras de teste pertencentes a outros lugares do mapa; e 75% dos testes indicando L0, para amostras pertencentes a L0.

Tabela 9 – Matriz de confusão* dos lugares para amostras de teste completas.

	L0	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12
L0	9	12	3	2	0	12	7	1	2	8	11	1	5
L1	0	12	0	0	0	0	0	0	0	0	0	0	0
L2	0	2	12	0	0	10	0	0	0	0	0	0	0
L3	4	12	11	9	2	12	7	0	6	5	10	2	1
L4	0	12	0	0	9	12	3	0	0	1	1	0	0
L5	0	0	0	0	0	12	0	0	0	0	0	0	0
L6	6	12	8	5	0	12	10	0	0	0	11	0	1
L7	10	12	12	11	4	12	12	10	11	12	10	10	12
L8	3	12	8	0	0	12	8	0	8	7	10	0	1
L9	0	12	2	0	0	12	0	0	0	12	4	0	0
L10	1	12	4	0	0	12	6	0	0	12	11	0	2
L11	7	12	11	11	0	12	10	0	12	8	12	8	2
L12	4	12	6	2	0	12	4	0	6	5	11	0	6
L13	0	12	7	2	0	12	8	0	5	5	5	1	7
L14	0	3	0	0	0	12	0	0	0	1	0	0	0
L15	9	12	12	10	7	12	12	4	12	12	12	10	12
L16	0	0	0	0	0	0	0	0	0	0	0	0	0
L17	0	0	0	0	0	1	0	0	0	0	0	0	0
L19	0	8	1	0	0	12	0	0	0	0	0	0	0
L20	0	0	0	0	0	3	0	0	0	0	0	0	0
L21	1	11	0	1	0	12	0	0	0	0	0	0	0
L22	0	0	0	0	0	4	0	0	0	0	0	0	0
L23	0	1	0	0	0	5	0	0	0	0	0	0	0
L24	0	1	0	0	0	10	0	0	0	0	0	0	0
L25	1	12	2	0	0	12	2	0	0	11	1	0	0
L26	0	12	2	0	0	12	6	0	0	1	0	0	0
L27	0	12	0	0	0	12	1	0	0	2	1	0	0

	L13	L14	L15	L16	L17	L19	L20	L21	L22	L23	L24	L25	L26	L27
L0	0	0	0	1	12	0	12	2	1	10	12	4	10	11
L1	0	0	0	0	0	0	5	0	0	0	0	0	0	0
L2	0	0	0	0	0	0	12	0	0	0	0	0	0	0
L3	0	0	0	1	12	11	12	5	6	12	11	4	4	12
L4	0	0	0	1	11	2	12	0	1	11	10	1	3	6
L5	0	0	0	0	0	0	1	0	0	0	0	0	0	0
L6	0	0	0	1	10	2	12	1	0	9	11	6	6	9
L7	0	2	4	11	12	12	12	12	12	12	12	12	12	12
L8	0	0	0	0	12	8	12	5	4	11	12	4	6	11
L9	0	0	0	0	8	3	12	0	0	6	4	0	0	4
L10	0	0	0	1	11	2	12	0	0	12	12	0	5	12
L11	0	1	1	5	12	10	12	9	4	12	12	10	9	12
L12	0	0	0	3	11	2	12	6	2	12	11	6	4	12
L13	8	2	1	6	12	1	12	2	10	12	12	9	9	11
L14	0	11	0	0	0	0	12	0	0	1	0	0	0	0
L15	3	9	10	10	12	11	12	11	12	12	12	12	12	12
L16	0	0	0	12	0	0	1	0	0	0	0	0	0	0
L17	0	0	0	0	12	0	0	0	0	0	0	0	0	0
L19	0	0	0	0	1	12	12	0	0	3	0	0	0	3
L20	0	0	0	0	0	0	12	0	0	0	0	0	0	0
L21	0	0	0	0	2	0	12	11	0	2	1	0	2	5
L22	0	0	0	0	0	0	10	0	12	0	0	0	0	0
L23	0	0	0	0	0	0	9	0	0	12	0	0	0	0
L24	0	0	0	0	0	0	8	0	0	0	12	0	0	0
L25	0	0	0	0	10	5	12	1	0	12	7	11	0	12
L26	0	0	0	0	8	0	12	0	0	12	6	1	10	8
L27	0	0	0	0	6	0	12	0	0	1	0	0	0	11

5.2.3 Análise dos Lugares com Desempenho Abaixo de Aceitável

Para os 37% dos lugares que obtiveram desempenho abaixo de aceitável com relação à análise das curvas ROC ($AUC \leq 0,7$), L0, L3, L6, L7, L8, L10, L11, L12, L13 e L15, confrontaram-se os resultados com os dados da matriz de confusão* da Tabela 9, com o intuito de descobrir o motivo de tal desempenho.

5.2.3.1 L7, L8 e L11

Ao observar imagens exemplo das amostras de caracterização e de teste dos lugares L7, L8 e L11, apresentadas na Figura 33, pode-se perceber que houve mudanças definitivas no ambiente, uma vez que partes desses lugares foram isoladas por conta de obras.

Essas mudanças são consideradas permanentes, pois não são apenas variações de objetos na cena, como ocorreu em todos os outros lugares do mapa, pois alteraram características importantes desses lugares, como destruição de um prédio e construção de outro, e também geraram oclusão de características de grande parte dos lugares. Além disso, tais mudanças se mantêm por um período de tempo prolongado.



Figura 33 – À esquerda temos imagens exemplo das amostras de caracterização e à direita de teste, correspondentes respectivamente de cima para baixo aos lugares L7, L8 e L11.

Logo, pode-se atribuir o desempenho ruim desses lugares às grandes variações no ambiente, as quais não podem ser consideradas como características dinâmicas dos lugares, pois os alteraram por tempo prolongado.

Para melhorar o desempenho de métodos de navegação que utilizem o modelo proposto neste trabalho, seria necessário atualizar as PMFs dos lugares apenas para mudanças significativas, tais como as descritas acima. Uma maneira de analisar de forma automática essa mudança no ambiente, seria monitorar o índice AUC ao longo do processo de navegação e estabelecer um limiar de AUC para desempenho mínimo esperado. Uma vez que o valor de AUC de um lugar fosse menor que o limiar definido pela curva ROC, a PMF desse lugar deveria ser redefinida.

5.2.3.2 L0, L3 e L10

Pela Tabela 8, que apresenta a quantidade de características extraídas por lugar do mapa, nota-se que os lugares L0 e L3 apresentam as maiores quantidades de características dentre os lugares do mapa, 91551 e 88423 respectivamente, se comparados ao lugar L20 que apresenta 19574 características e desempenho de classificação superior.

Como já foi observado na Seção 5.1 e confirmado na Seção 5.2.2, lugares com grande quantidade de características tendem a apresentar mal desempenho. Dessa forma, conclui-se que a influência da quantidade de características extraídas é um dos principais fatores para o desempenho ruim desses lugares.

No caso do lugar L10, apesar deste não ser um dos lugares com maior quantidade de características extraídas, este fator pode ser considerado o motivo do seu desempenho não tão bom ($AUC = 0,6835$), pois este lugar apresenta 60915 características e este valor está próximo de 66146, que é a média dos lugares que apresentaram desempenho abaixo do aceitável.

5.2.3.3 L12 e L13

Ao observar a disposição dos lugares no mapa na Figura 29, a primeira suspeita do desempenho ruim dos lugares L12 e L13 foi a proximidade destes e suas semelhanças. No entanto, ao analisar a matriz de confusão* na Tabela 9 pôde-se concluir que essa suspeita não foi confirmada, pois as mostras de L12 são mais localizadas como L1, L5, L10, L17, L20, L23, L24 e L27 (pelo menos 92% das amostras) do que como L13 (0%), e as amostras de L13 são mais localizadas como L1, L5, L17, L20, L23, L24 e L27 (pelo menos 92% das amostras) do que como L12 (58%).

Observando a Tabela 8, que apresenta a quantidade de características extraídas por lugar do mapa, nota-se que os lugares L12 e L13 apresentam grandes quantidades de características, 59812 e 52816 respectivamente, se comparados ao lugar L20 que apresenta 19574 e desempenho superior. Dessa forma, conclui-se que a influência da quantidade de características extraídas é um dos principais fatores para o desempenho ruim desses lugares.

Além disso, ao observar a Tabela 7 é possível perceber que o horário de captura das imagens de caracterização é favorável a situações de baixa iluminação, por ser fim de tarde. Após analisar as imagens exemplo de caracterização 30m e 30n, nota-se realmente um baixo nível de iluminação, principalmente nas imagens de L13, fato que pode influenciar negativamente o processo de caracterização dos lugares por dificultar a extração de características e a determinação de correspondências.

5.2.3.4 L6 e L15

Para os lugares L6 e L15, pode-se perceber problemas parecidos com os problemas encontrados nos lugares L12 e L13. Pela Tabela 8, que apresenta a quantidade de características extraídas por lugar do mapa, nota-se que tanto L6 quanto L15 apresentam grandes quantidades de características, 45741 e 58312 respectivamente. Como já discutido anteriormente, esse fato contribui para queda de desempenho de classificação.

Além disso, pela Tabela 7 percebe-se que o horário de captura das imagens de caracterização tanto de L6 quanto de L15, também é favorável a situações de baixa iluminação, fato comprovado pelas imagens 30g e 30p. Como já dito anteriormente, esse fato pode afetar negativamente o processo de caracterização dos lugares e consequentemente o método de localização.

Para comprovar a influência negativa da iluminação no processo de caracterização dos lugares, e consequentemente no método de localização, será apresentada uma análise mais abrangente sobre este fator na Seção 5.3.

5.3 Influência da Iluminação

Foi levada em consideração a variabilidade de níveis de iluminação⁵ no processo de escolha do período de aquisição das amostras, ou seja, tentou-se capturar situações que pudessem levar o modelo a apresentar um desempenho ruim. Caso estas situações fossem identificadas, e seus efeitos negativos confirmados, seria possível definir estratégias que pudessem melhorar o desempenho do modelo ao evitar ou até mesmo torná-lo robusto a tais situações.

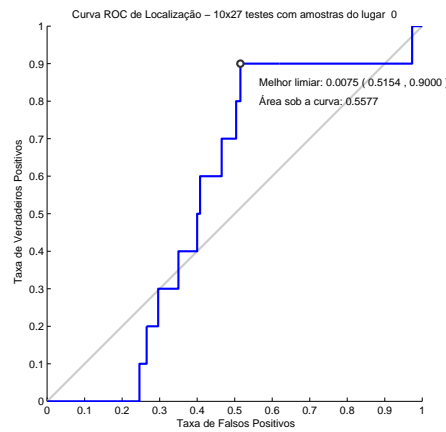
Em testes preliminares percebeu-se que para amostras de teste capturadas sob níveis de iluminação parecidos com os da captura de suas respectivas amostras de caracterização, o desempenho mínimo do modelo foi excelente ($AUC > 0,8$). Por isso, a captura da maioria das imagens foi feita em diferentes períodos do dia para as imagens de teste se comparadas às imagens de caracterização.

⁵ Não foram feitas medições com sensores de iluminação para determinar tais níveis, no entanto é possível perceber ou qualificá-los visualmente pelas imagens exemplo das amostras.

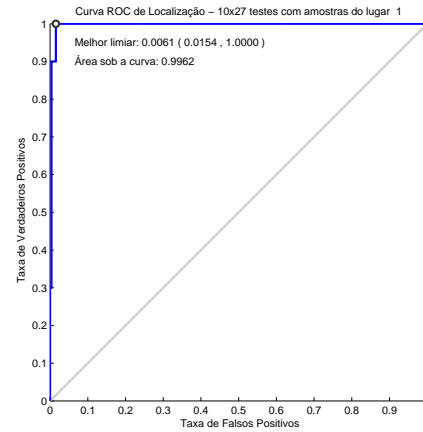
Pela Tabela 7 percebe-se que na maioria dos casos tem-se a seguinte regra: se as imagens de caracterização foram capturadas próximo ao meio-dia, as imagens de teste foram capturadas próximo às seis horas da tarde e vice-versa.

Logo, pela Tabela 7 pode-se observar que os lugares candidatos a apresentar problemas por conta da iluminação são L1, L4, L5, L6, L9, L11, L12, L13, L14, L15, L16, L17, L19, L20, L21 e L27. No entanto, alguns desses lugares já apresentaram desempenho pelo menos aceitável, segundo a Figura 32. Assim, selecionou-se os lugares L4, L6, L9, L11, L12, L13 e L15 para inverter as amostras de teste com amostras de caracterização, com o objetivo de analisar o efeito da melhoria do nível de iluminação nas amostras de caracterização.

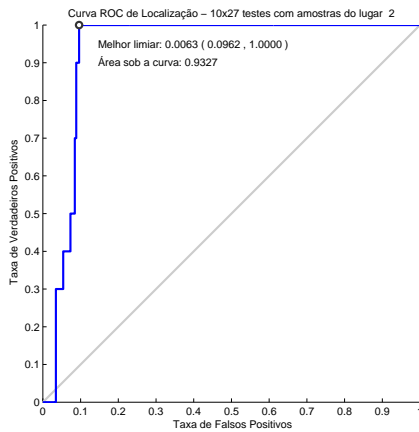
Após realizar o processo descrito na Seção 4.1, foram geradas curvas ROC para 10 amostras de caracterização dos lugares, e utilizando 10 amostras completas para teste. Com isso, foram obtidas as curvas ROC apresentadas na Figura 34.



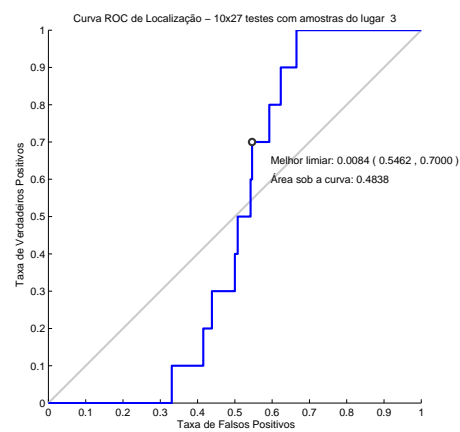
(a) Lugar L0.



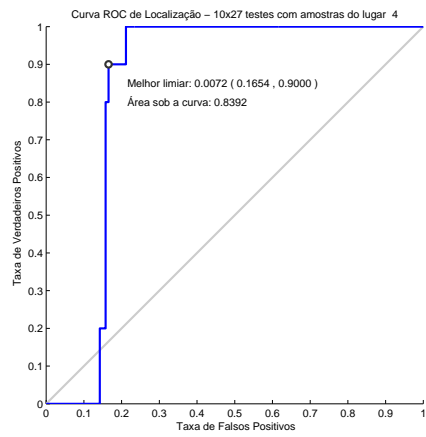
(b) Lugar L1.



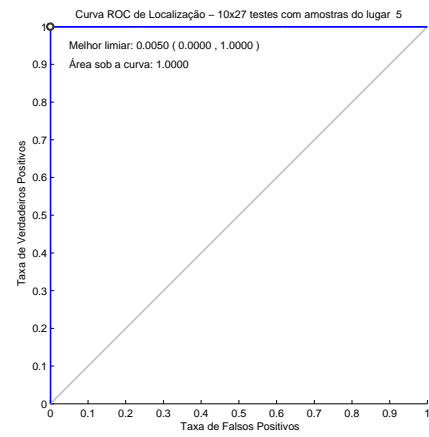
(c) Lugar L2.



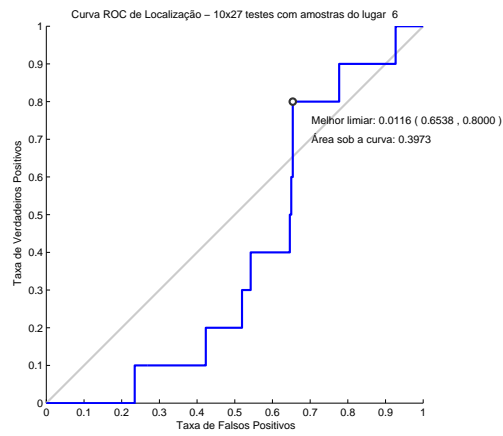
(d) Lugar L3.



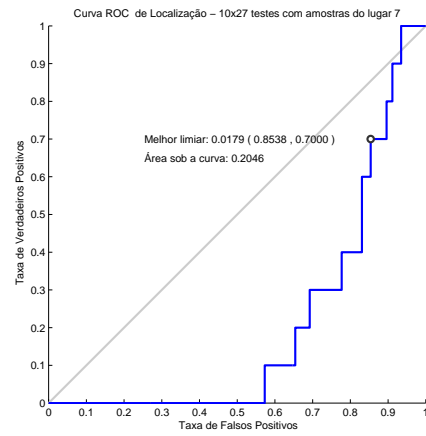
(e) Lugar L4.



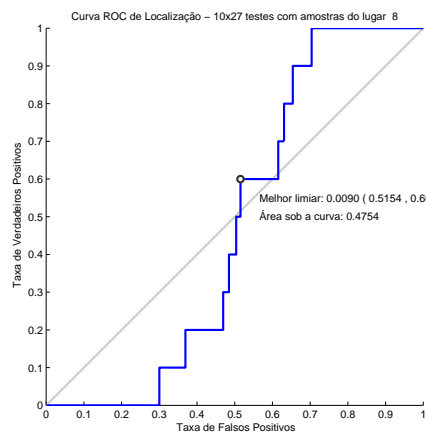
(f) Lugar L5.



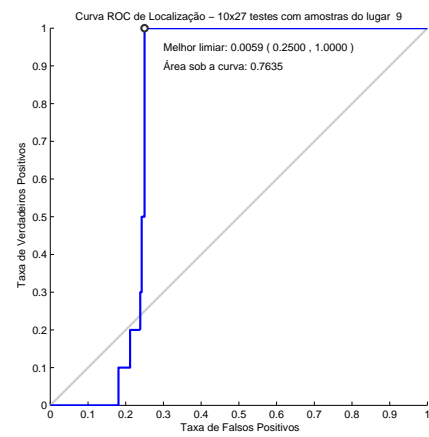
(g) Lugar L6.



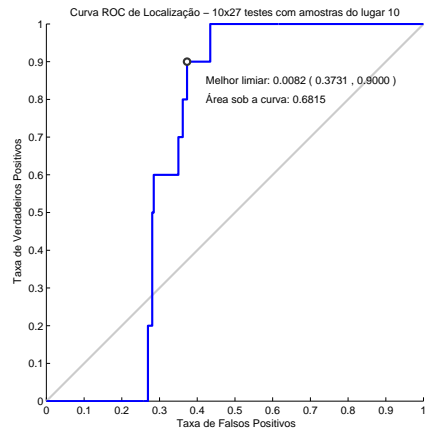
(h) Lugar L7.



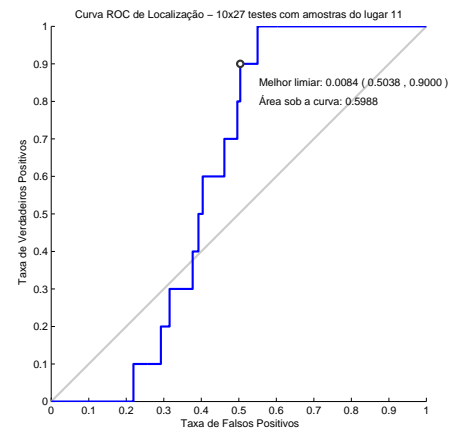
(i) Lugar L8.



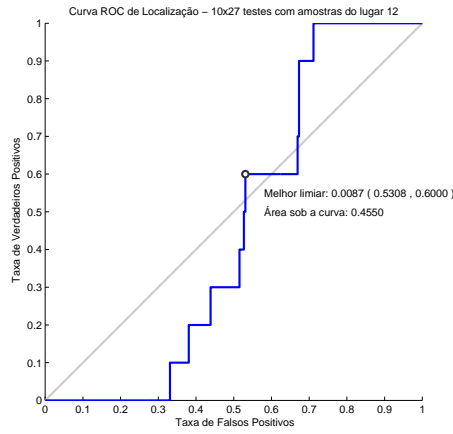
(j) Lugar L9.



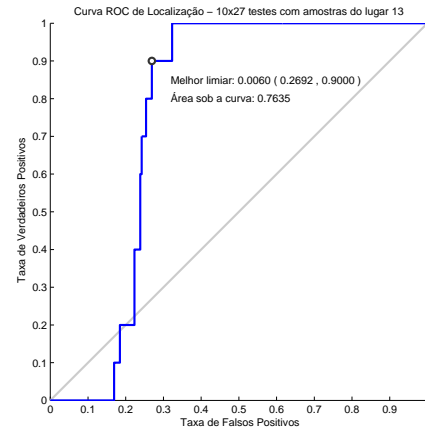
(k) Lugar L10.



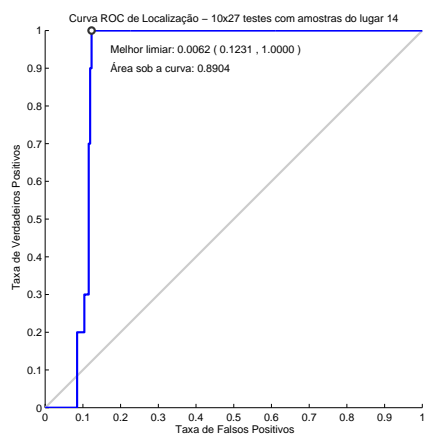
(l) Lugar L11.



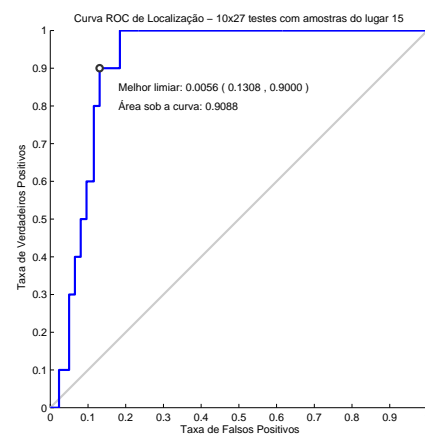
(m) Lugar L12.



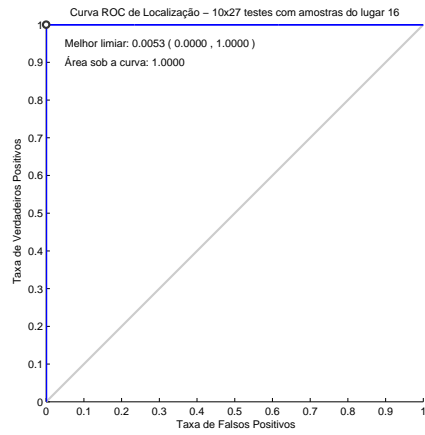
(n) Lugar L13.



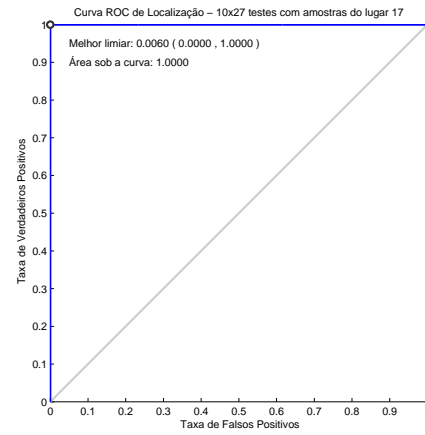
(o) Lugar L14.



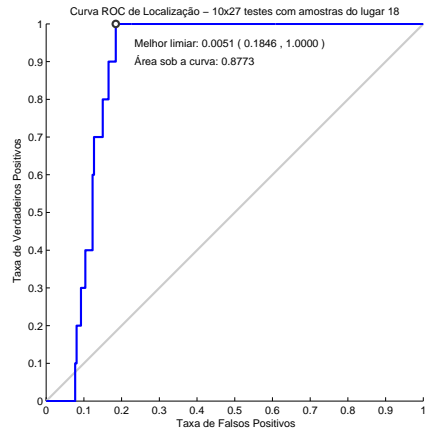
(p) Lugar L15.



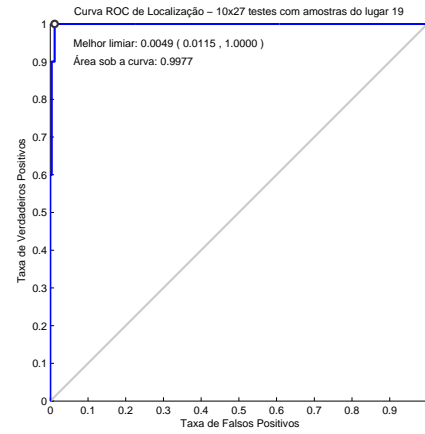
(q) Lugar L16.



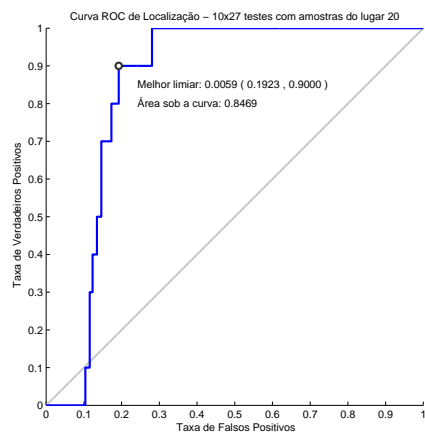
(r) Lugar L17.



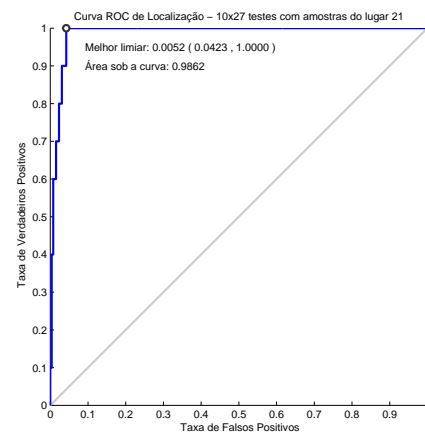
(s) Lugar L19.



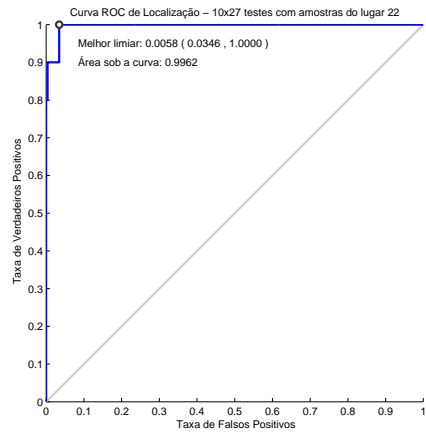
(t) Lugar L20.



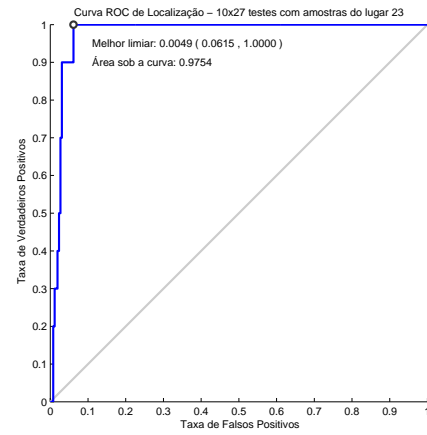
(u) Lugar L21.



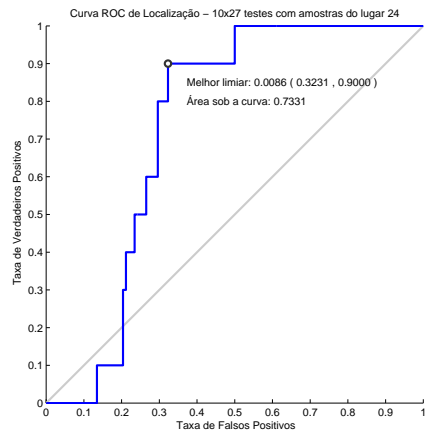
(v) Lugar L22.



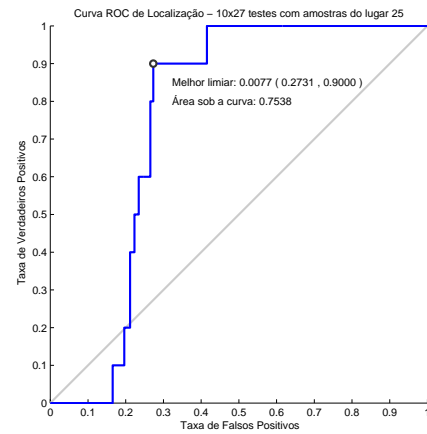
(w) Lugar L23.



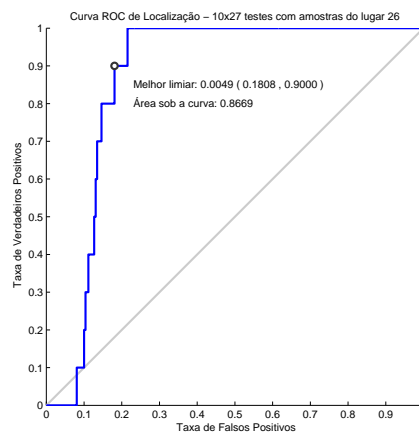
(x) Lugar L24.



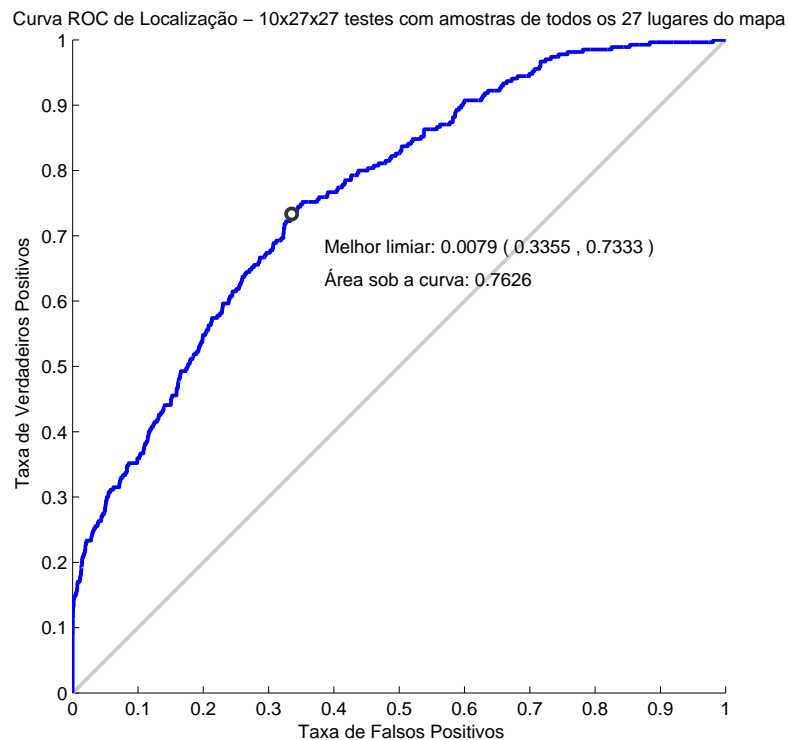
(y) Lugar L25.



(z) Lugar L26.



(aa) Lugar L27.



(ab) Mapa.

Figura 34 – Curvas ROC dos lugares e do mapa, geradas por 10 amostras completas de teste por lugar, totalizando 270 testes por lugar e 7290 para o mapa.

O primeiro fato que pode ser observado é que houve uma pequena melhora no desempenho global do modelo, pois a AUC do mapa passou de 0,7366 para 0,7626. Contudo, dos lugares que tiveram suas imagens de caracterização invertidas com as imagens de teste, apenas os lugares L4, L11, L13 e L15 obtiveram melhoria no seu desempenho, que sem a inversão apresentavam respectivamente valores de AUC iguais a 0,7767; 0,3649; 0,4712 e 0,2014; segundo a Figura 32. Já após a inversão, passaram a 0,8392; 0,5988; 0,7635 e 0,9088; segundo a Figura 34.

Para que possa ser efetuada uma análise da influência do nível de iluminação sobre o modelo, também deve ser levada em consideração a quantidade de características extraídas dos lugares. Se o aumento do nível de iluminação realmente influenciar positivamente o desempenho do lugar, este fator pode ser influenciado negativamente pelo aumento relativo do número de características, e como resultante pode-se piorar o desempenho dos lugares ao melhorar o nível de iluminação.

Analisando a Tabela 10, percebe-se que L4, L11 e L13 obtiveram melhoria no desempenho, mesmo tendo aumentado a quantidade de características por conta do

aumento do nível de iluminação. Logo, conclui-se que a melhoria do nível de iluminação nas imagens de caracterização pode melhorar tanto o desempenho global quanto local do modelo.

Entretanto, o desempenho dos lugares afetados por esse fator pode piorar se o efeito negativo do aumento das características for superior ao efeito positivo da melhoria do nível de iluminação, como ocorreu nos lugares L6, L9 e L12, através da Figura 34.

Um fato que chama atenção ao analisar a Tabela 10, é a diminuição da quantidade de características do lugar L15 após a inversão das imagens de teste com as de caracterização. Ao observar as Figuras 30p e 31p é possível perceber qualitativamente, que o nível de iluminação na imagem de teste, Figura 31p, é menor que na imagem de caracterização, Figura 30p.

Tabela 10 – Quantidade de características extraídas dos lugares do mapa antes e após a inversão das imagens de teste com as imagens de caracterização.

Identificação	Nº de Características	Nº de Características Após Inversão	Diferença Percentual
L4	53223	69682	31%
L6	45741	72358	58%
L9	67129	68222	2%
L11	72389	82363	14%
L12	59812	83407	39%
L13	52816	65634	24%
L15	58312	40144	-31%

Logo, o processo de inversão das imagens não alcançou o efeito desejado para o lugar L15, que seria melhorar o nível de iluminação das imagens de caracterização. Todavia, comprova-se mais uma vez o efeito da quantidade de características sobre o modelo, pois ao reduzir a quantidade de características em 31%, o valor de AUC da curva ROC do lugar L15 passou de 0,2014 para 0,9088.

Tal efeito também pode ser comprovado pela matriz de confusão* apresentada na Tabela 11, que se comparada com a matriz de confusão* antes da inversão, apresentada na Tabela 9, mostra uma diminuição de localizações erradas para amostras de L15.

Tabela 11 – Matriz de confusão* dos lugares para amostras de teste completas após a inversão de imagens de teste com imagens de caracterização de alguns lugares do mapa.

	L0	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12
L0	9	10	2	2	3	10	3	2	4	2	10	9	3
L1	0	10	0	0	0	0	0	0	0	0	0	0	0
L2	0	2	10	0	0	8	0	0	0	0	0	0	0
L3	1	10	9	6	5	10	4	0	1	3	8	9	2
L4	0	8	0	0	9	10	0	0	0	0	0	0	0
L5	0	0	0	0	0	10	0	0	0	0	0	0	0
L6	1	10	9	0	9	10	6	0	9	1	6	7	1
L7	5	10	10	7	10	10	5	7	10	5	8	9	7
L8	3	10	7	0	5	10	0	0	6	0	7	6	4
L9	0	10	0	0	0	10	0	0	0	10	0	0	0
L10	3	10	4	0	4	10	0	0	0	0	9	0	1
L11	6	10	10	0	10	10	0	0	1	0	6	9	0
L12	0	10	8	0	2	10	0	0	4	0	3	0	6
L13	0	10	0	0	0	10	0	0	0	0	0	0	0
L14	0	1	0	0	0	10	0	0	0	0	0	0	0
L15	0	7	0	0	0	10	0	0	0	0	0	0	0
L16	0	0	0	0	0	0	0	0	0	0	0	0	0
L17	0	0	0	0	0	1	0	0	0	0	0	0	0
L19	0	6	1	0	1	10	0	0	0	0	0	5	0
L20	0	0	0	0	0	3	0	0	0	0	0	0	0
L21	0	10	0	0	0	10	0	0	0	0	0	2	0
L22	0	0	0	0	0	3	0	0	0	0	0	0	0
L23	0	0	0	0	0	4	0	0	0	0	0	0	0
L24	0	0	0	0	0	8	0	0	0	0	0	0	0
L25	0	10	2	0	1	10	0	0	0	0	1	0	0
L26	0	10	0	0	0	10	0	0	0	0	0	0	0
L27	1	10	0	0	1	10	0	0	0	0	1	1	0

	L13	L14	L15	L16	L17	L19	L20	L21	L21	L23	L24	L25	L26	L27
L0	1	0	10	2	10	1	10	3	1	8	10	3	9	10
L1	0	0	0	0	0	0	5	0	0	0	0	0	0	0
L2	0	0	5	0	0	0	10	0	0	0	0	0	0	0
L3	0	0	10	1	9	8	10	4	5	10	9	4	0	10
L4	0	0	8	0	8	0	10	0	0	0	0	0	0	0
L5	0	0	0	0	0	0	1	0	0	0	0	0	0	0
L6	8	1	10	10	10	2	10	2	4	10	10	10	10	10
L7	8	2	10	10	10	10	10	8	9	10	10	10	10	10
L8	2	0	10	0	10	7	10	4	3	9	10	3	5	9
L9	0	0	10	0	6	0	10	0	0	7	1	0	1	10
L10	3	0	10	0	9	1	10	0	0	10	10	0	4	10
L11	1	0	10	0	10	7	10	4	4	10	10	1	3	10
L12	9	0	10	1	9	4	10	3	8	10	9	9	10	9
L13	9	0	10	0	8	0	10	0	0	10	5	1	1	5
L14	0	9	10	0	0	0	10	0	0	1	0	0	0	0
L15	0	0	9	0	3	0	10	0	0	0	0	0	5	0
L16	0	0	0	10	0	0	1	0	0	0	0	0	0	0
L17	0	0	0	0	10	0	0	0	0	0	0	0	0	0
L19	0	0	8	0	1	10	10	0	0	3	0	0	0	3
L20	0	0	0	0	0	0	10	0	0	0	0	0	0	0
L21	0	0	9	0	2	0	10	8	0	2	0	0	2	3
L22	0	0	0	0	0	0	8	0	10	0	0	0	0	0
L23	0	0	0	0	0	0	6	0	0	10	0	0	0	0
L24	0	0	1	0	0	0	7	0	0	0	10	0	0	0
L25	0	0	10	0	9	4	10	1	0	10	7	9	0	10
L26	3	0	10	0	7	0	10	0	0	10	5	1	9	6
L27	0	0	7	0	5	0	10	0	0	1	0	0	0	9

5.4 Utilização de Amostras de Teste Incompletas

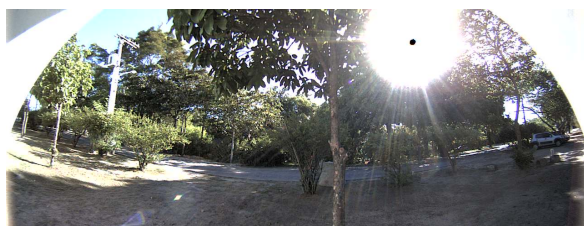
Como já discutido no Capítulo 3, a utilização de amostras completas para caracterização dos lugares se faz necessária para que seja possível captar o máximo de informação. Contudo, dependendo do grau de semelhança entre os lugares do mapa, pode ser possível utilizar amostras de teste incompletas e ainda assim alcançar um desempenho aceitável para o modelo.

Esta abordagem é justificada pela diminuição de imagens de teste, e consequentemente pela redução do tempo necessário para efetuar a localização de uma amostra, tanto pelo método local proposto na Equação 3.9, quanto por um método global. Quanto menos imagens para compor uma amostra, menor será o tempo para extrair as características e menor será a quantidade de características a buscar correspondências na PMF dos lugares.

Para comprovar a suposição de que a utilização de amostras incompletas pode resultar em um desempenho pelo menos aceitável, foi construído um mapa com os mesmos 28 lugares do mapa da Figura 29, agora incluindo o lugar L18⁶, com as mesmas amostras de caracterização da Seção 5.2 e mantendo a mesma metodologia. Vale ressaltar que o mapa foi construído com amostras completas e apenas as mostras de teste eram incompletas.

Para isso, foram capturadas 16 novas amostras de teste, representadas pelas imagens exemplo da Figura 35, agora compostas por apenas duas imagens. Em todos os lugares foi garantida uma distância mínima de aproximadamente 4 metros entre o ponto de captura de uma imagem de caracterização e o ponto de captura de sua respectiva imagem de teste.

⁶ O lugar L18 pôde ser incluído pois no período de captura das amostras de teste deste lugar, o mesmo ainda não tinha sido isolado.



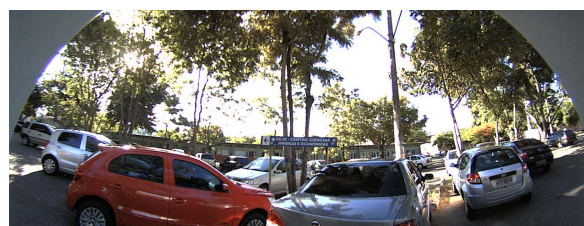
(a) Lugar L0.



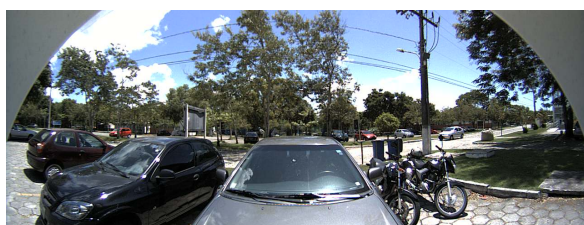
(b) Lugar L1.



(c) Lugar L2.



(d) Lugar L3.



(e) Lugar L4.



(f) Lugar L5.



(g) Lugar L6.



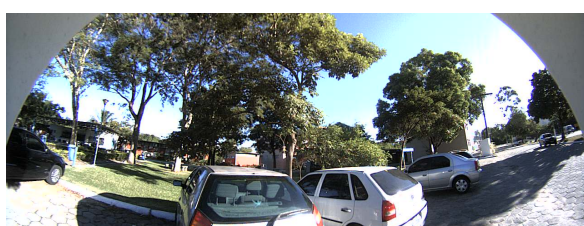
(h) Lugar L7.



(i) Lugar L8.



(j) Lugar L9.



(k) Lugar L10.



(l) Lugar L11.



(m) Lugar L12.



(n) Lugar L13.



(o) Lugar L14.



(p) Lugar L15.



(q) Lugar L16.



(r) Lugar L17.



(s) Lugar L18.



(t) Lugar L19.



(u) Lugar L20.



(v) Lugar L21.



(w) Lugar L22.



(x) Lugar L23.

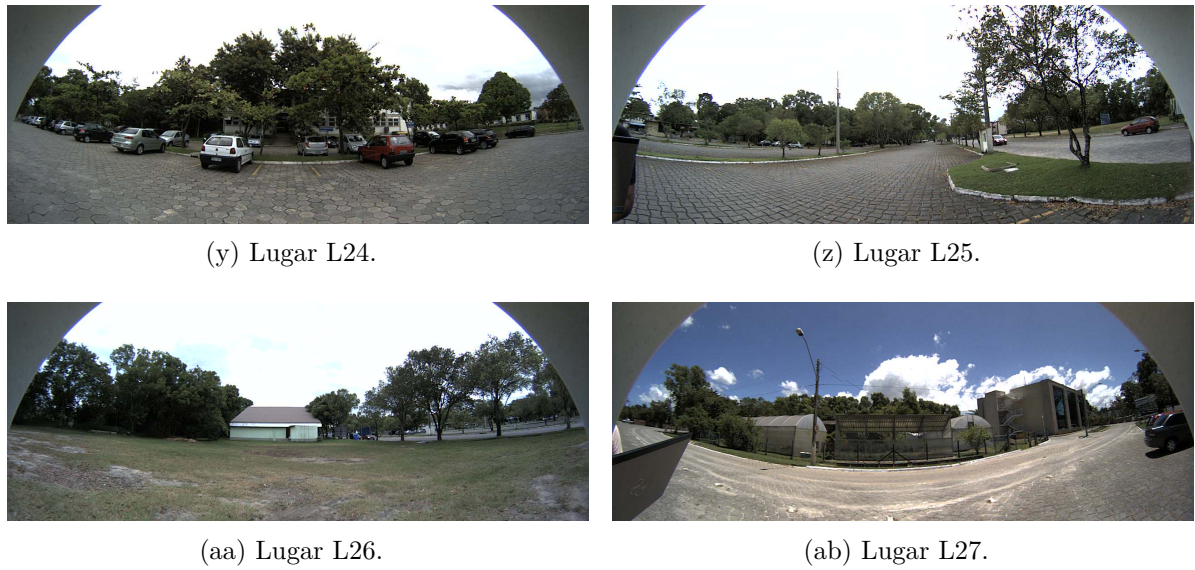


Figura 35 – Imagens exemplo das amostras de teste incompletas dos lugares do mapa.

Pela Tabela 12, percebe-se que os intervalos de tempo entre a captura das imagens de caracterização e das imagens de teste foram bem menores, exceto nos lugares L16, L17 e L18, se comparados aos intervalos apresentados na Tabela 7 do experimento da Seção 5.2.

Tal fato reflete diretamente no desempenho do modelo ao utilizar amostras de teste incompletas, como pode ser observado na Figura 36, que mostra um aumento no índice AUC da curva ROC do mapa, em relação ao experimento da Figura 32, que utilizou amostras completas.

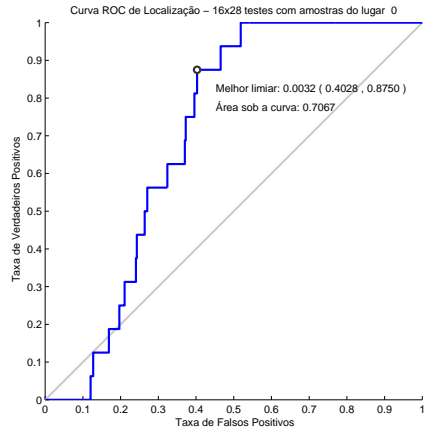
Comparando as curvas ROC das Figuras 36 e 32, observa-se que no experimento com amostras completas 37% dos lugares apresentaram desempenho abaixo do aceitável, e ao utilizar amostras incompletas esse número caiu para 29%. Todavia, em 63% dos lugares houve redução no índice AUC, e essa redução apresentou média de 13%.

Dessa forma, conclui-se que ao utilizar amostras incompletas para realizar a localização (amostras de teste), pode-se alcançar desempenho global aceitável, se não houver alterações permanentes no ambiente. No entanto, o desempenho local pode ser severamente prejudicado se comparado à utilização de amostras completas, visto que houve redução no índice AUC em 63% dos lugares, e a maior redução foi de 75% no lugar L14.

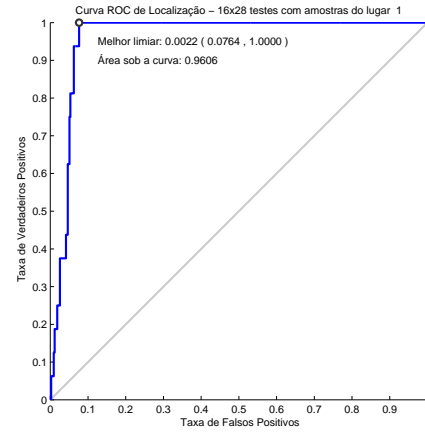
Além disso, pôde-se confirmar a redução no tempo médio de localização das amostras no melhor caso, visto que no teste com amostras completas o tempo médio de localização no melhor caso foi 26,082 segundos e com amostras incompletas foi de 12,283 segundos, ou seja, redução de 53%. A análise detalhada sobre questões temporais a respeito da construção do mapa e da localização de amostras será apresentada no Capítulo 6.

Tabela 12 – Períodos de captura das amostras de teste e de caracterização, utilizadas no experimento com amostras incompletas. Na segunda e terceira colunas temos o dia, a hora e os minutos de início e fim da captura, e na quarta coluna temos o intervalo de tempo mínimo (em meses) entre a captura das imagens de caracterização e das imagens de teste.

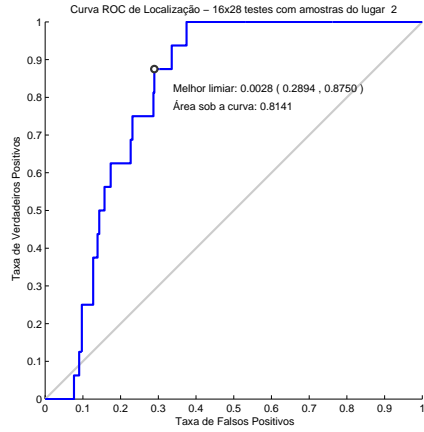
	Captura das Amostras de Caracterização	Captura das Amostras de Teste	Intervalo Mínimo (meses)
L0	17.01.2012 / 11:00 - 11:05	29.02.2012 / 15:55 - 15:59	2
L1	28.01.2012 / 17:22 - 17:25	01.02.2012 / 12:21 - 12:24	0
L2	17.01.2012 / 11:10 - 11:14	29.02.2012 / 16:09 - 16:13	1
L3	17.01.2012 / 11:38 - 11:40	29.02.2012 / 16:20 - 16:24	1
L4	12.01.2012 / 17:09 - 17:12	28.02.2012 / 11:43 - 11:46	1
L5	28.01.2012 / 16:53 - 16:56	01.02.2012 / 13:02 - 13:04	0
L6	28.01.2012 / 17:13 - 17:16	01.02.2012 / 12:50 - 17:52	0
L7	12.01.2012 / 11:52 - 11:57	30.01.2012 / 14:57 - 15:03	0
L8	17.01.2012 / 11:21 - 11:24	02.03.2012 / 16:23 - 16:25	1
L9	12.01.2012 / 17:48 - 17:50	28.02.2012 / 11:37 - 11:41	1
L10	17.01.2012 / 11:28 - 11:32	02.03.2012 / 16:30 - 16:33	1
L11	12.01.2012 / 16:58 - 17:02	06.03.2012 / 12:19 - 12:23	1
L12	12.01.2012 / 17:18 - 17:23	08.03.2012 / 11:42 - 11:44	1
L13	12.01.2012 / 17:28 - 17:30	06.03.2012 / 12:39 - 12:41	1
L14	12.01.2012 / 17:35 - 17:39	06.03.2012 / 12:50 - 12:53	1
L15	16.01.2012 / 18:04 - 18:10	08.03.2012 / 12:00 - 12:02	1
L16	28.01.2012 / 16:35 - 16:39	25.01.2011 - 22.03.2011 / 10:00 - 18:00	10
L17	28.01.2012 / 16:46 - 16:49	22.03.2011 / 16:17 - 16:20	10
L18	14.03.2011 - 22.03.2011 / 12:00 - 17:00	25.01.2011 - 14.03.2011 / 12:00	13
L19	16.01.2012 / 17:55 - 17:59	08.03.2012 / 11:50 - 11:53	1
L20	28.01.2012 / 16:59 - 17:02	01.02.2012 / 12:55 - 12:57	0
L21	28.01.2012 / 17:07 - 17:10	01.02.2012 / 12:44 - 12:47	0
L22	12.01.2012 / 12:39 - 12:42	01.02.2012 / 12:39 - 12:41	0
L23	12.01.2012 / 12:32 - 12:35	01.02.2012 / 12:32 - 12:35	0
L24	12.01.2012 / 12:21 - 12:24	30.01.2012 / 15:47 - 15:50	0
L25	12.01.2012 / 12:06 - 12:08	30.01.2012 / 15:24 - 15:27	0
L26	12.01.2012 / 12:14 - 12:17	30.01.2012 / 15:33 - 15:38	0
L27	16.01.2012 / 18:17 - 18:19	28.02.2012 / 11:30 - 11:33	1



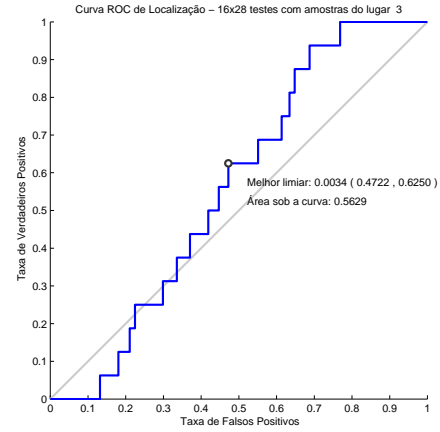
(a) Lugar L0.



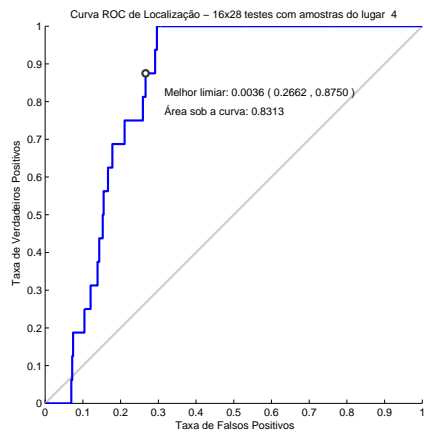
(b) Lugar L1.



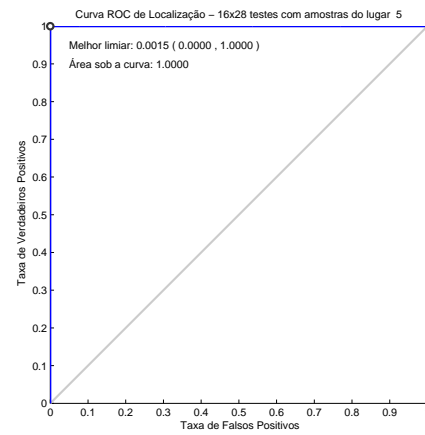
(c) Lugar L2.



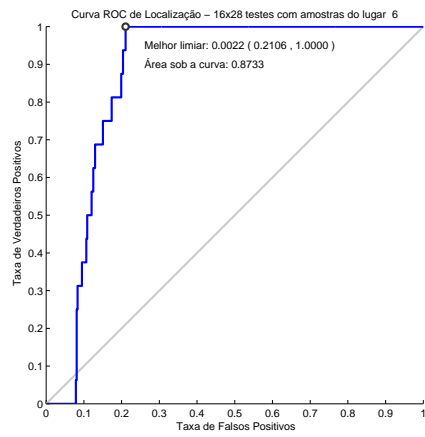
(d) Lugar L3.



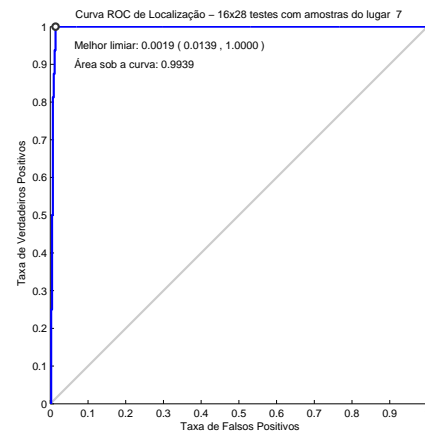
(e) Lugar L4.



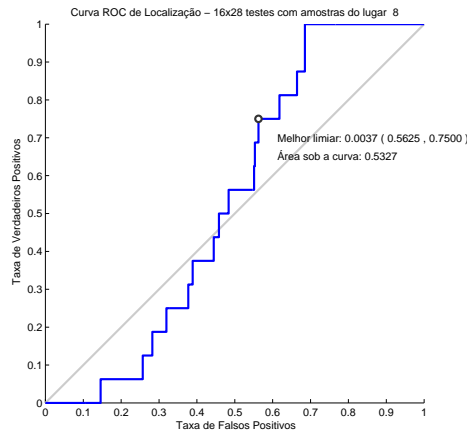
(f) Lugar L5.



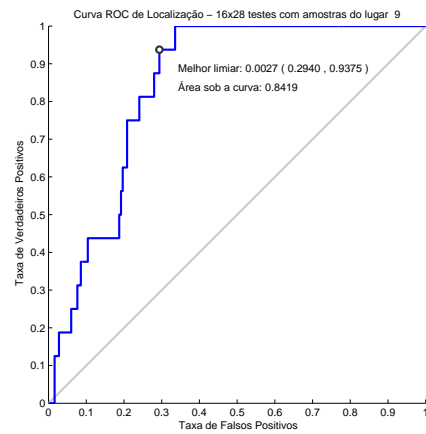
(g) Lugar L6.



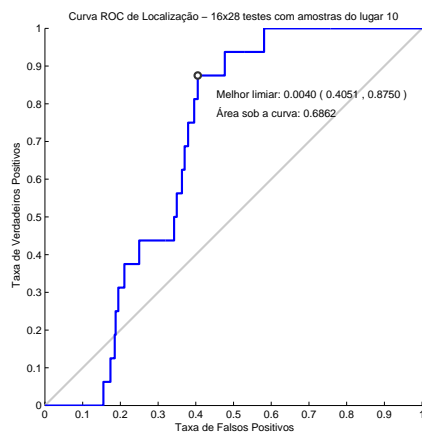
(h) Lugar L7.



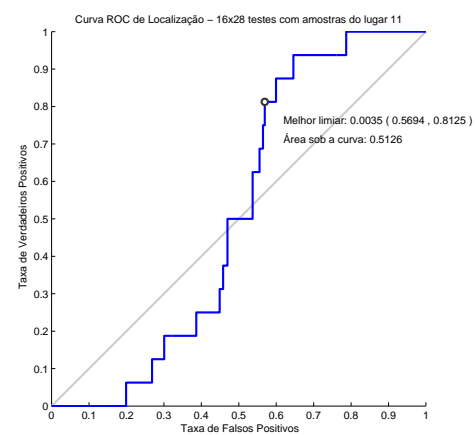
(i) Lugar L8.



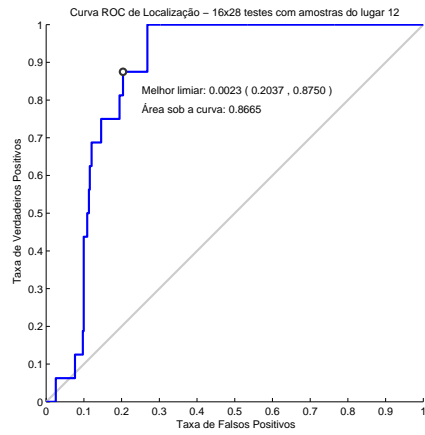
(j) Lugar L9.



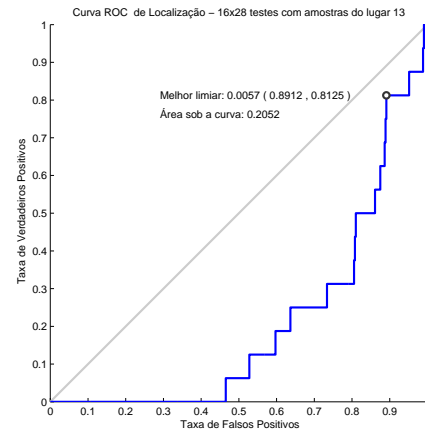
(k) Lugar L10.



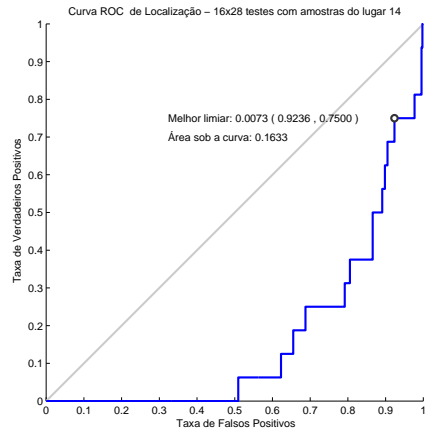
(l) Lugar L11.



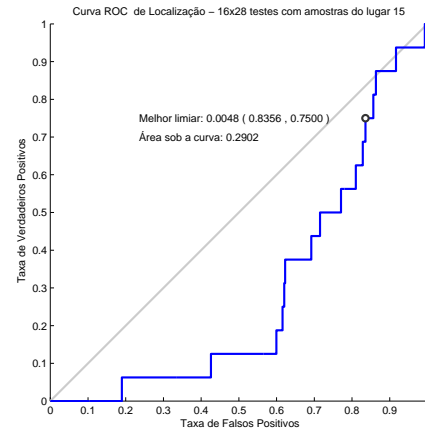
(m) Lugar L12.



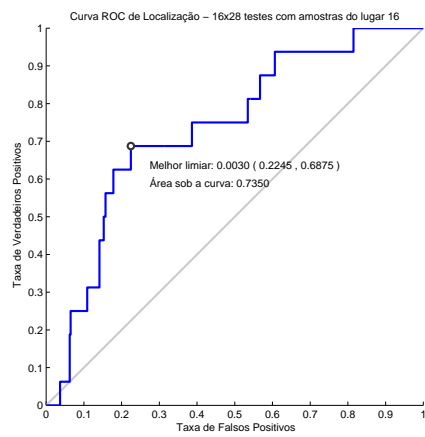
(n) Lugar L13.



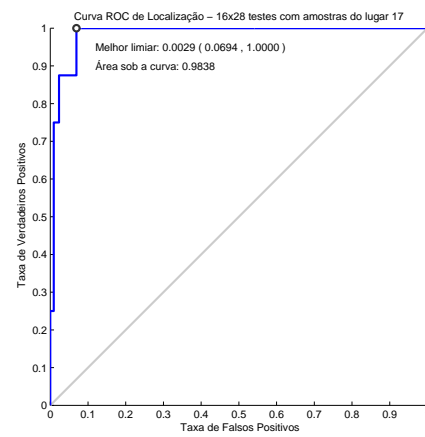
(o) Lugar L14.



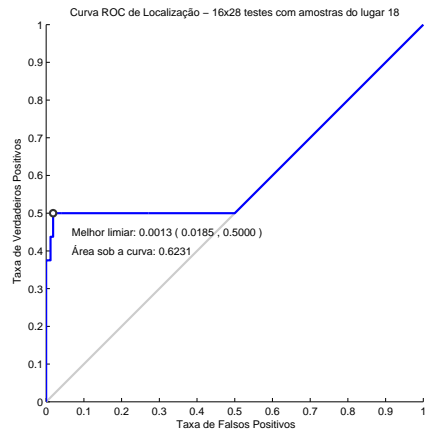
(p) Lugar L15.



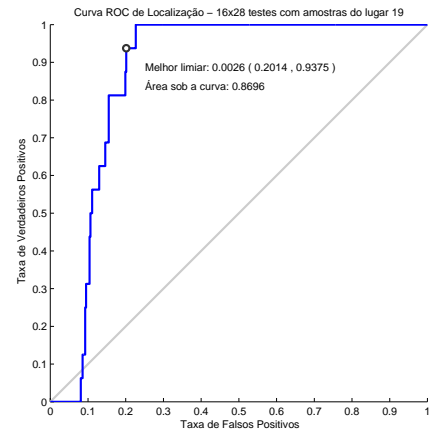
(q) Lugar L16.



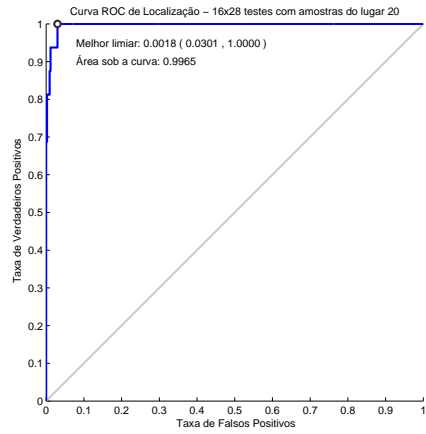
(r) Lugar L17.



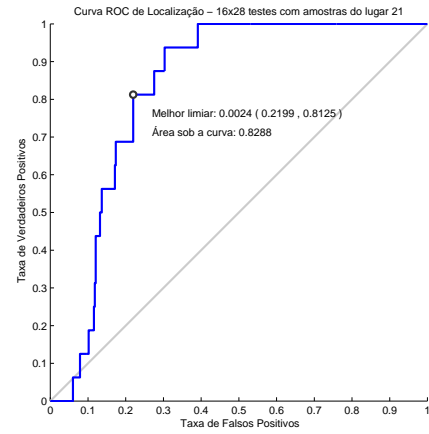
(s) Lugar L18.



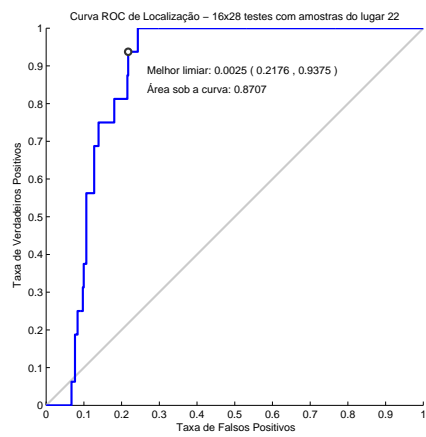
(t) Lugar L19.



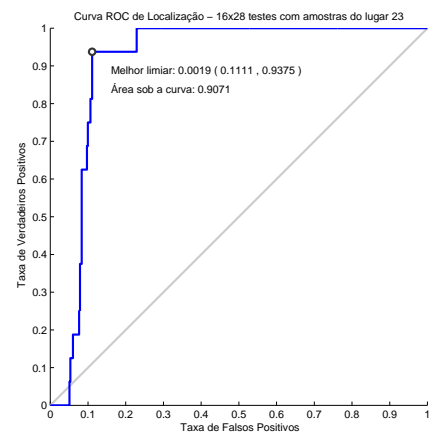
(u) Lugar L20.



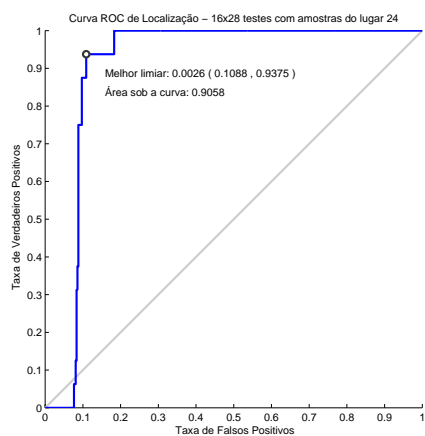
(v) Lugar L21.



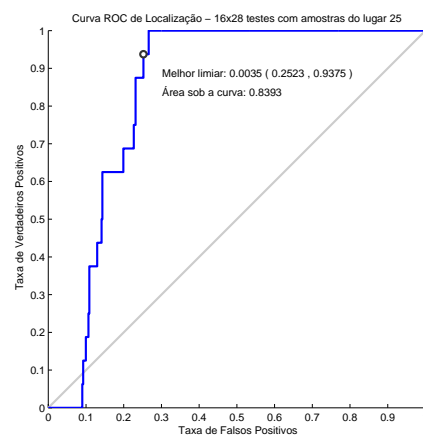
(w) Lugar L22.



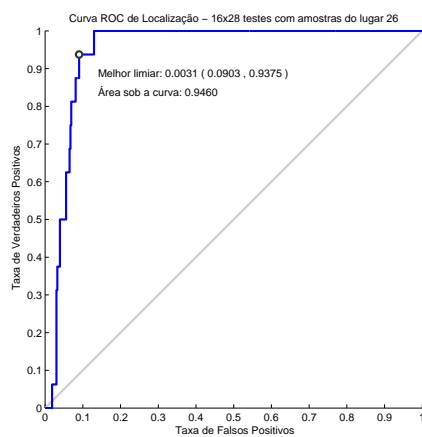
(x) Lugar L23.



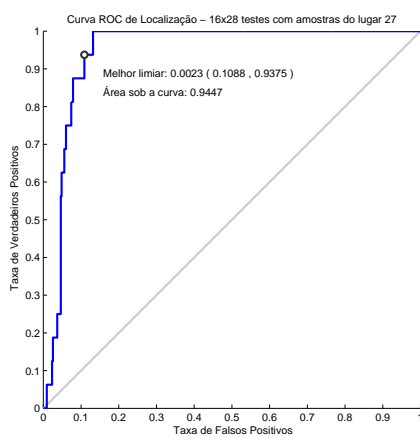
(y) Lugar L24.



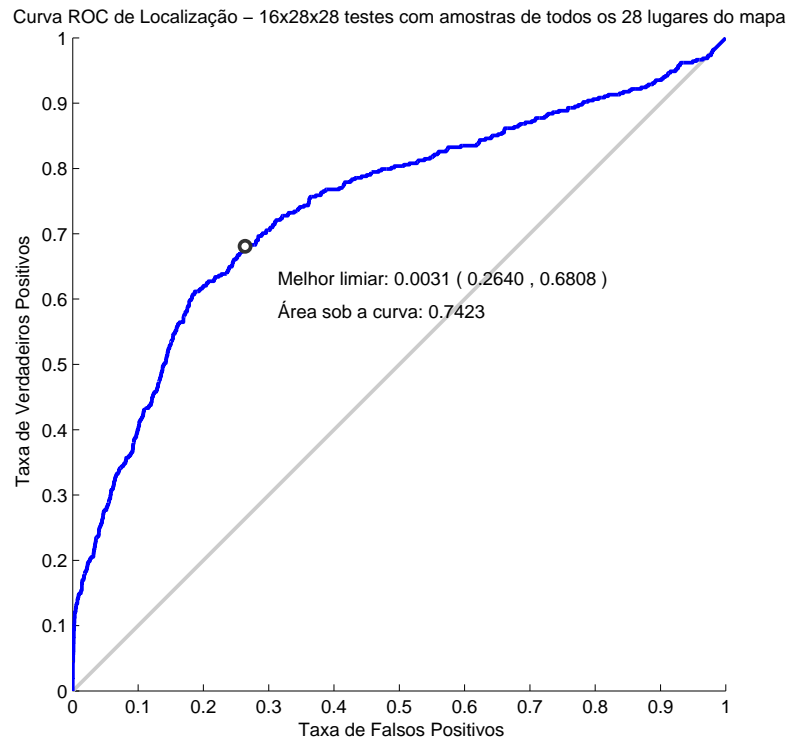
(z) Lugar L25.



(aa) Lugar L26.



(ab) Lugar L27.



(ac) Mapa.

Figura 36 – Curvas ROC dos lugares e do mapa, geradas por 16 amostras incompletas (2 imagens apenas) de teste por lugar, totalizando 448 testes por lugar e 12544 para o mapa.

5.5 Descarte de Características da PMF dos Lugares

Através dos resultados obtidos e de suas respectivas análises, é possível concluir que o modelo proposto pode ser utilizado para realizar localização em ambientes externos. No entanto, para que esta localização seja efetuada de forma *online*, como proposto no objetivo deste trabalho, é preciso determinar o tempo gasto para efetuar a localização de uma amostra no melhor caso, para que seja possível estimar o tempo gasto pelo método de localização no pior caso.

Por isso, como já apresentado anteriormente, foi calculado o tempo médio de 26,082 segundos⁷ no melhor caso para calcular a probabilidade de amostras completas (4 imagens) e de 12,283 segundos no melhor caso para amostras incompletas (2 imagens),

⁷ Apesar desse tempo, e de todos os outros apresentados nesta seção, terem sido alcançados pela utilização de um processador com 4 núcleos, pode-se garantir que tal desempenho foi devido prioritariamente à utilização de um disco rígido de 7200 RPM, e não à quantidade de núcleos do processador, como será discutido no Capítulo 6.

respectivamente para os experimentos das Figuras 32 e 36.

Como pode-se concluir, mesmo utilizando amostras incompletas o tempo de cálculo das probabilidades é muito alto para que este modelo seja utilizado em métodos de navegação *online*. Pois, no experimento com 27 lugares e amostras de teste completas da Figura 32, o tempo médio de localização das amostras foi de 10,2495 minutos⁸ no pior caso de localização, no qual é necessário calcular a probabilidade da amostra pertencer à todos os lugares do mapa para efetuar a localização. Logo, para um mapa com n lugares este tempo seria de aproximadamente $O(26,082 * n)$ segundos.

Então, identificou-se a necessidade de reduzir o tempo de localização, e para isso observou-se que a quantidade de características das PMFs dos lugares poderiam ser reduzidas, por um algoritmo que eliminasse destas algumas características indesejadas. Assim, embora o tempo de construção do mapa não fosse reduzido, pois todas as características presentes na imagens já teriam sido extraídas, o tempo de localização das amostras poderia ser reduzido de forma significativa, pois com menos características nas PMFs dos lugares, menor seria o tempo gasto para determinação de correspondências entre as características das amostras e das PMFs.

Assim, foi elaborado um algoritmo que após a definição de uma PMF, descarta características que tenham sido observadas um número de vezes menor que um limiar definido como entrada do algoritmo. Então, para visualizar o efeito desse algoritmo no valor da probabilidade das amostras de teste, montou-se um experimento para o qual foram capturadas 5 amostras completas para 10 pontos, definidos de 4 em 4 metros ao longo de 36 metros em torno do lugar L23⁹. Em seguida, foram calculadas as médias das probabilidades das 5 amostras de cada ponto para três situações, considerando os limiares 1, 8 e 10 para o algoritmo de descarte das características da PMF.

Pelo gráfico apresentado na Figura 37, é possível perceber que as probabilidades das amostras de teste diminuem com a redução da quantidade de características da PMF de L23 (limiares 10 e 8). Além disso, percebe-se que a utilização apenas das características fortes¹⁰, definida pela utilização do limiar 10, gera valores de probabilidade próximos aos valores calculados pela utilização do limiar 8, o qual considera características que ocorreram pelo menos 8 vezes no lugar L23.

Segundo a Tabela 13, observa-se que o aumento do limiar no algoritmo de descarte diminui a quantidade de características na PMF do lugar, diminui o tempo médio gasto para

⁸ Neste tempo de localização computou-se a extração de características SURF de cada amostra de teste apenas uma vez.

⁹ A escolha do lugar L23 foi feita simplesmente por conta da facilidade em capturar imagens sob várias posições neste local.

¹⁰ Dado que foram utilizadas m amostras completas no processo de definição de uma PMF, considera-se que uma característica forte é aquela observada m vezes no conjunto de imagens de caracterização do lugar.

calcular a probabilidade de uma amostra de teste pertencer ao lugar, e consequentemente diminui o tempo de localização das amostras de teste, pois pelo método de localização proposto na Equação 3.9, basta comparar o valor da probabilidade da amostra com o valor de probabilidade de referência do lugar para definir se a amostra pertence ou não ao lugar em questão.

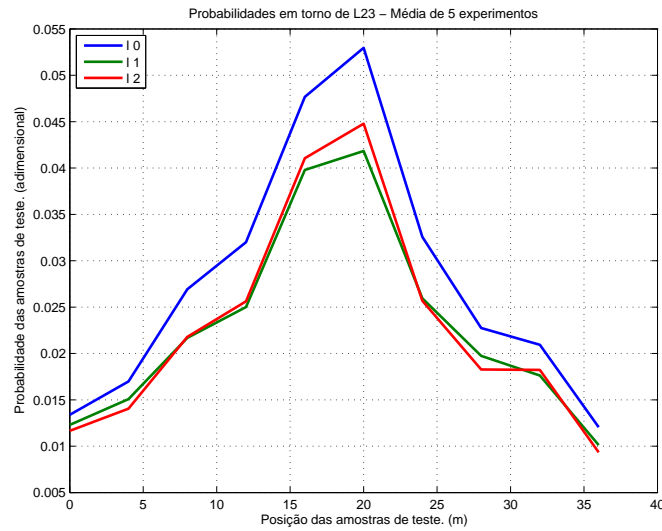


Figura 37 – Comparação das probabilidades das amostras em torno do lugar L23, para diferentes limiares utilizados para descartar características da PMF de L23. Em azul a curva l0 de limiar 1, em verde a curva l1 de limiar 8, e em vermelho a curva l2 de limiar 10.

Tabela 13 – Número de características da PMF de L23, e tempo de médio de cálculo das probabilidades das amostras de teste pertencerem ao lugar L23 para os limiares 1, 8 e 10, do algoritmo de descarte de características.

Limiar da PMF	Nº de Características	Tempo Médio do Cálculo da Probabilidade (s)
1	36508	24,805
8	16969	13,306
10	8374	8,498

Contudo, para comprovar a hipótese de que a redução de características de todos os lugares do mapa pode não gerar grandes perdas de desempenho do modelo, repetiram-se os experimentos das Seções 5.2.1 e 5.4 que utilizaram 10 amostras de caracterização, e respectivamente 12 e 16 amostras de teste para mapas compostos respectivamente por 27 e 28 lugares, porém utilizando o algoritmo de descarte para três limiares, 1, 8 e 10. Nas Figuras 38 e 39 as três curvas ROC dos mapas são comparadas.

Percebe-se que o desempenho do modelo confirma o resultado apresentado na Figura 37, pois o melhor desempenho do modelo é alcançado para a utilização de todas as características da PMF dos lugares para ambos os experimentos, e a utilização de características fortes apresenta desempenho similar à utilização de características que ocorreram menos vezes que a quantidade de amostras de caracterização, 10 neste mapa.

Pela Tabela 14, confirma-se também a redução no tempo de localização das amostras com o aumento do limiar do algoritmo de descarte, pois observa-se uma redução de 77% no tempo médio de localização para amostras completas no melhor caso, e 73% para amostras incompletas, com a utilização do limiar 10.

Logo, considera-se que o processo de descarte de características é uma boa abordagem para reduzir o tempo de localização das amostras e consequentemente obter um método de localização que possa ser utilizado por algoritmos de navegação *online*.

Entretanto, a elaboração de um algoritmo que selecione um pequeno conjunto de características consideradas representativas do lugar, e não apenas características fortes, é um bom exemplo de meta futura para este trabalho. Dessa forma, será possível obter tempos de localização similares aos da Tabela 14, reduzindo menos o desempenho do modelo se comparado aos resultados das Figuras 38 e 39.

Tabela 14 – Tempo de médio de localização das amostras de teste no melhor caso, tanto para o experimento com amostras completas quanto incompletas, para os limiares 1, 8 e 10, do algoritmo de descarte de características da PMF dos lugares do mapa.

Limiar da PMF	Amostras Completas	Amostras Incompletas
	Tempo Médio de Localização no Melhor Caso (s)	
1	26,082	12,283
8	10,081	5,542
10	6,129	3,361

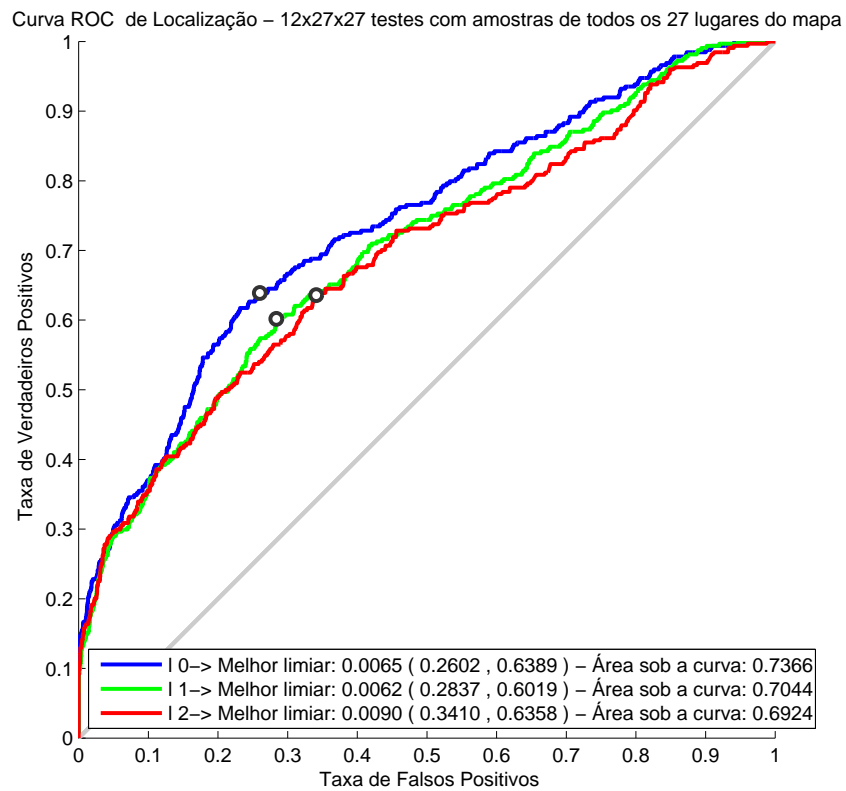


Figura 38 – Comparação das curvas ROC do mapa com amostras teste completas, para três limiares de descarte das características das PMFs. Em azul a curva l0 de limiar 1, em verde a curva l1 de limiar 8, e em vermelho a curva l2 de limiar 10.

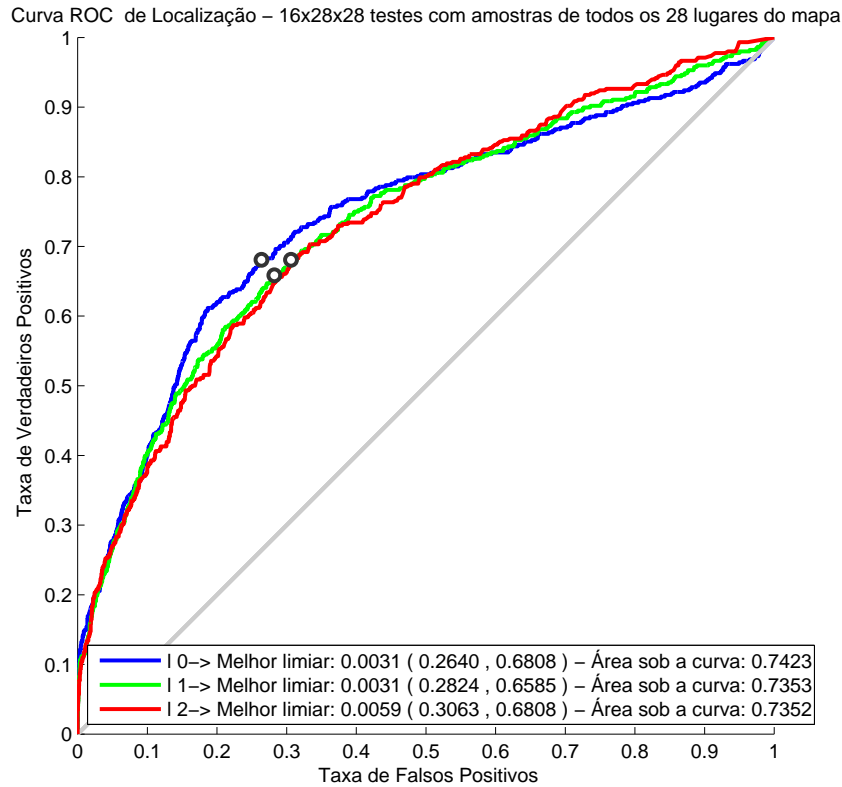


Figura 39 – Comparação das curvas ROC do mapa com amostras de teste incompletas, para três limiares de descarte das características das PMFs. Em azul a curva l0 de limiar 1, em verde a curva l1 de limiar 8, e em vermelho a curva l2 de limiar 10.

5.6 Comportamento em Regiões de Transição entre Lugares

Embora as conclusões sobre o desempenho do modelo tenham sido favoráveis à sua utilização, faz-se necessária uma análise sobre o comportamento do mesmo em regiões do ambiente que não tenham sido mapeadas, como apresentado na Seção 3.1.1.

Logo, para analisar o comportamento das probabilidades de amostras de teste nestas regiões, escolheram-se dois lugares que apresentam visada direta entre si. O objetivo dessa escolha foi checar a existência de relação entre a probabilidade das amostras e a distância destas à posição de caracterização dos lugares, pois comprovando esta hipótese seria possível mostrar a possibilidade de localização mesmo em lugares não mapeados.

Por isso, elaborou-se um experimento ao longo de 96 metros em torno dos lugares L23 e L22, escolhidos por causa de suas similaridades e por apresentarem visada direta entre si. Tal experimento consistiu em calcular a probabilidade de amostras de teste

completas em relação tanto a L23 quanto a L22, de 4 em 4 metros, sendo que este processo foi repetido 5 vezes. Vale ressaltar que a caracterização desses lugares foi feita 4 meses antes da realização do experimento. Assim, a partir desse experimento foram geradas as Figuras 40, 41 e 42.

Observando a Figura 42, percebe-se que há um crescimento não monotônico da probabilidade, quando as amostras se aproximam das posições de caracterização dos lugares. Dessa forma, considere o método de localização proposto na Equação 3.9, com valores de referência 0,0052 para L23 e 0,0048 para L22, definidos no experimento da Seção 5.2.1.

Dessa forma, ao longo dos 96 metros do experimento as amostras poderiam ser localizadas tanto pertencentes ao lugar L23 quanto ao lugar L22. Logo, é necessário definir um critério para reduzir a chance de erro de localização do método proposto na Equação 3.9. Um critério a ser agregado ao método é a análise da derivada da função de probabilidade pela posição das amostras, como apresentado a seguir

$$Crit_{L_i} = \frac{d(Pr_{L_i})}{d_x}. \quad (5.1)$$

Para definir a localização das amostras utilizando a função gerada pela Equação 5.1 é preciso encontrar um ponto de máximo nesta função. Dessa forma, é necessário continuar calculando as probabilidades das amostras em relação ao local indicado pelo método de localização, até ser possível encontrar um ponto de máximo na função gerada pela Equação 5.1.

Para o caso da Figura 42, considere que as amostras estão sendo capturas no sentido de L23 para L22 e que o método de localização indicou L22 como sendo a localização das amostras. O critério da Equação 5.1 eliminaria o erro de localização até a posição 72 metros, e portanto, o método de localização continuaria procurando outro lugar candidato enquanto permanecesse nesta região. Caso a localização indicada pelo método fosse L23, o critério de localização confirmaria este resultado pouco depois da posição 20 metros.

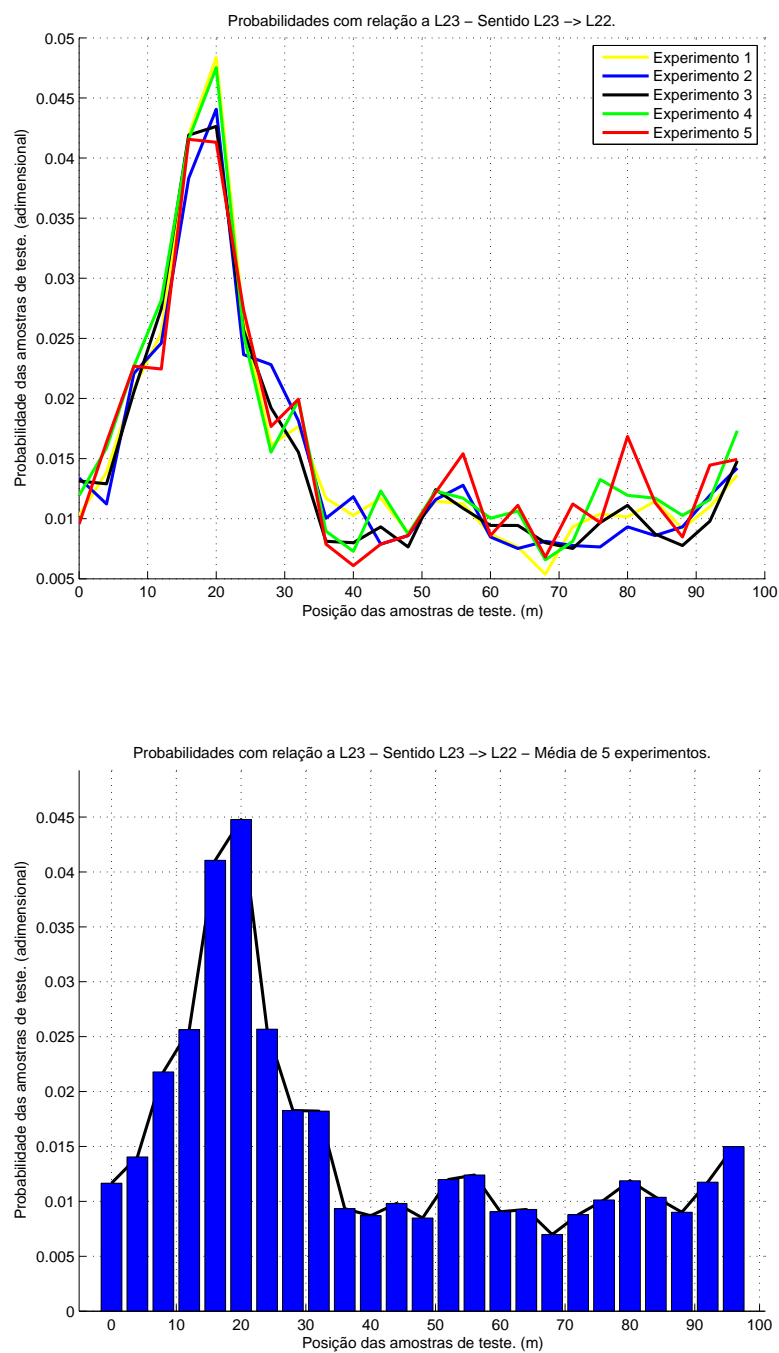


Figura 40 – Probabilidades das amostras em relação a L23, acima temos os 5 experimentos e abaixo a média destes experimentos.

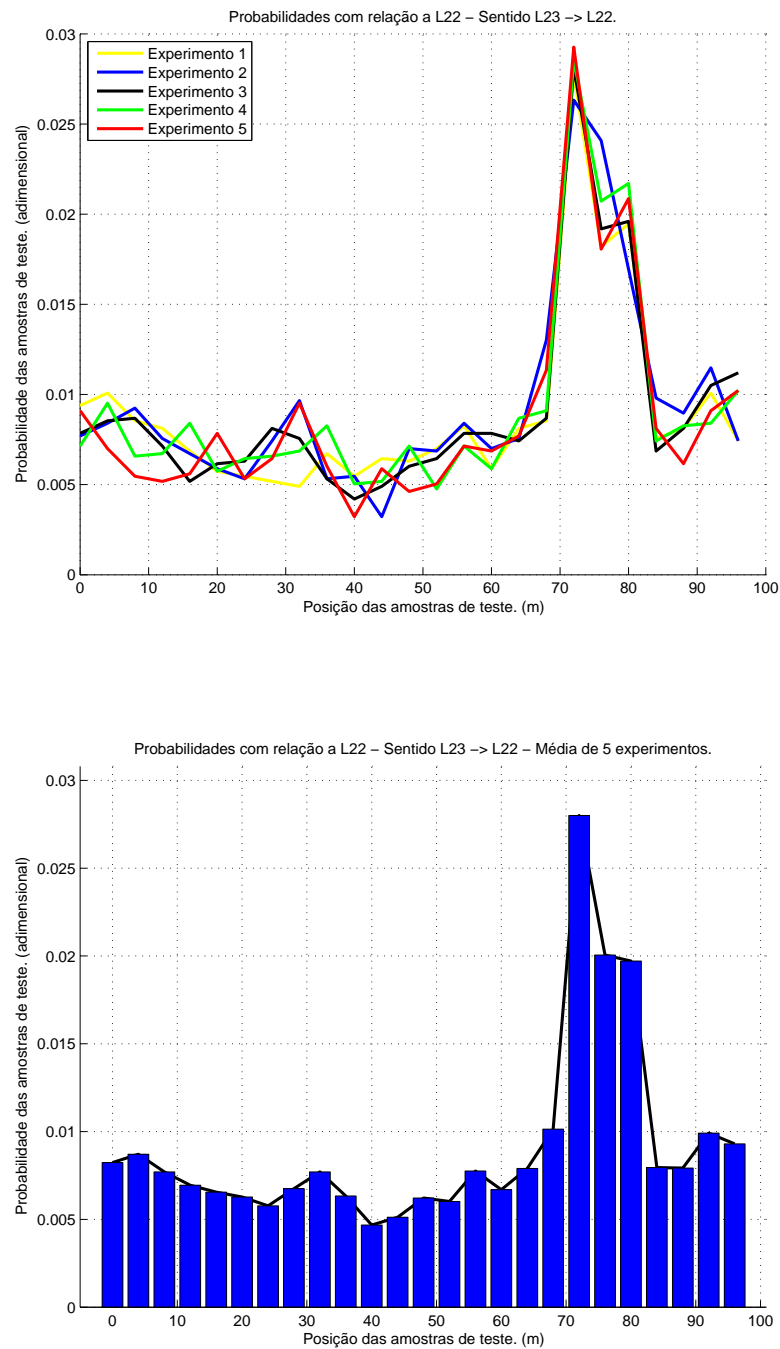


Figura 41 – Probabilidades das amostras em relação a L22, acima temos os 5 experimentos e abaixo a média destes experimentos.

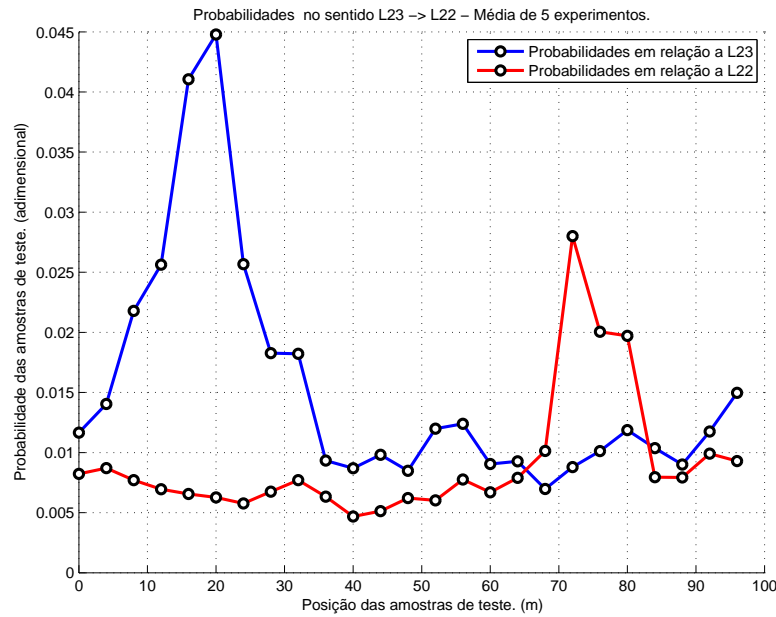


Figura 42 – Média das probabilidades dos 5 experimentos, em relação a L23 e a L22.

5.7 O Problema de Reconhecimento de Objetos

Para avaliar o modelo de caracterização de cenas apresentado no Capítulo 3, aplicado ao problema de reconhecimento de objetos, foram efetuados testes com um conjunto de 4 objetos caracterizados por 3 amostras completas, e 5 amostras de teste completas, utilizando imagens com resolução de 640x480. Neste caso, considerou-se que uma amostra completa é composta por duas imagens das superfícies de maior área.

Na escolha dos objetos optou-se por objetos que apresentassem características similares para submeter o modelo a uma situação extrema, semelhante a de visada direta das seções anteriores. Além disso, para capturar as imagens de teste utilizou-se um fundo diferente do utilizado sobre as amostras caracterização, com o intuito de mostrar que o modelo consegue diferenciar os objetos do fundo. Nas Figuras 43 e 44 são apresentadas imagens exemplo, das amostras de caracterização e teste dos objetos.

Após realizar o processo descrito na Seção 4.1, foram geradas curvas ROC para as 3 amostras de caracterização dos objetos, e foram utilizadas 5 amostras completas para teste. Os resultados são apresentados na Figura 45.

Nota-se que o desempenho global do classificador foi aceitável e um objeto apresentou desempenho superior, Figura 45c. Este resultado, é confirmado pela matriz de confusão* apresentada na Tabela 15, que mostra por exemplo, o objeto com maior desempenho de classificação, faca, sendo menos confundido com os outros objetos que os demais.

Além disso, o experimento foi repetido utilizando o algoritmo de descarte de características e os resultados obtidos, apresentados na Figura 46 e na Tabela 16, confirmam o resultado já apresentado para a aplicação do modelo ao problema de localização: o aumento do limiar no algoritmo de descarte gera perda de desempenho de classificação.

No entanto, utilização de somente características fortes através do limiar 3, que causou perda de desempenho de classificação redução do tempo de reconhecimento no melhor caso, não se faz necessária para esse classificador, pois o tempo de reconhecimento com a utilização de todas as características das PMFs já foi aceitável por estar abaixo de 1 segundo.

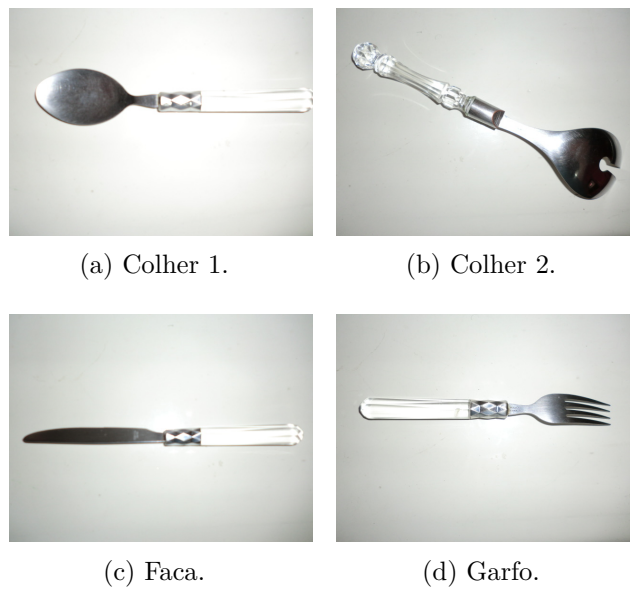


Figura 43 – Imagens exemplo das amostras de caracterização dos objetos.

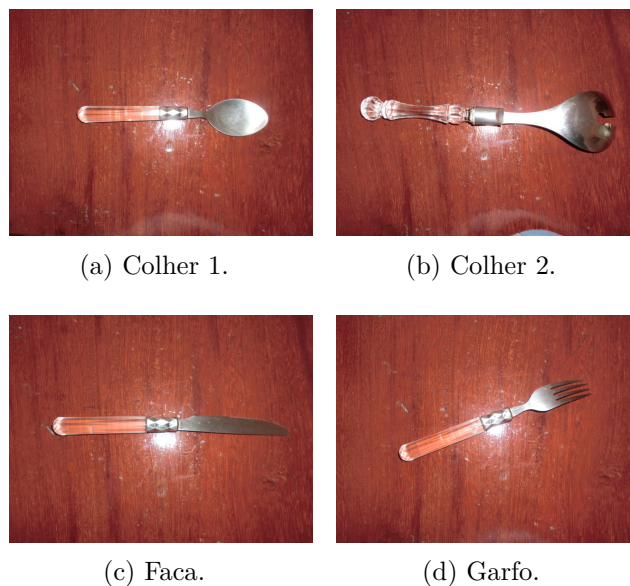
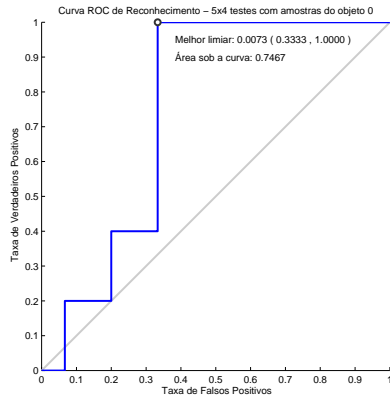
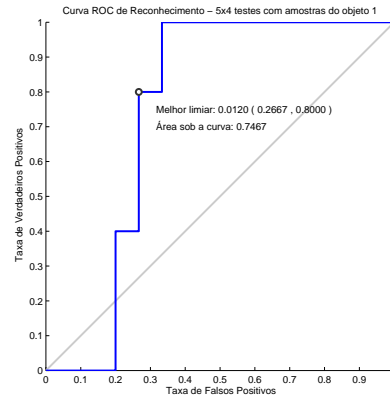


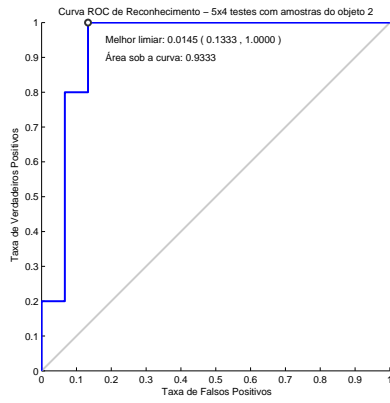
Figura 44 – Imagens exemplo das amostras de teste dos objetos.



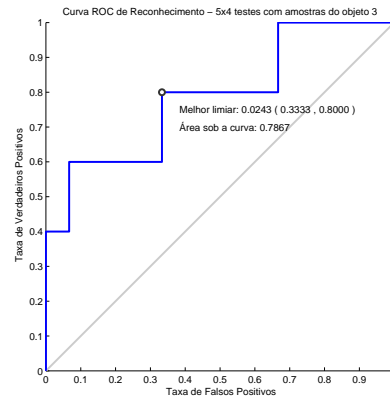
(a) Colher 1.



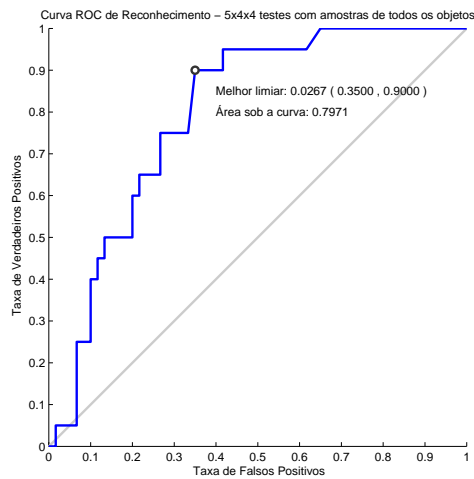
(b) Colher 2.



(c) Faca.



(d) Garfo.



(e) Classificador.

Figura 45 – Curvas ROC dos objetos e do classificador, geradas por 5 amostras completas (2 imagens apenas) de teste por objeto, totalizando 20 testes por objeto e 80 para o classificador.

Tabela 15 – Matriz de confusão* dos objetos, para 5 amostras de teste completas.

	Colher 1	Colher 2	Faca	Garfo
Colher 1	5	0	1	4
Colher 2	2	4	2	1
Faca	1	1	5	1
Garfo	3	1	1	4

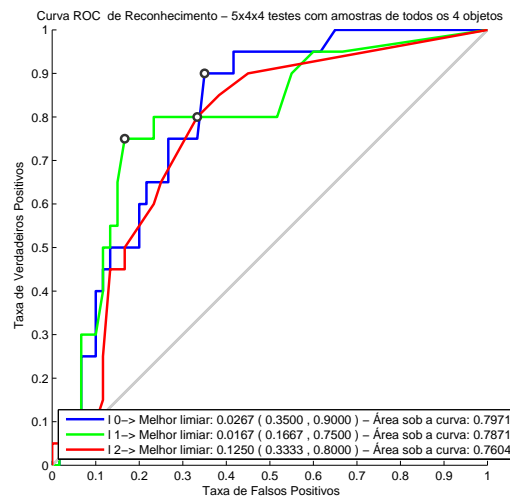


Figura 46 – Comparação entre as curvas ROC do classificador. Em azul a curva l0 de limiar 1, em verde a curva l1 de limiar 2, e em vermelho a curva l2 de limiar 3.

Tabela 16 – Tempo de médio de classificação das amostras de teste no melhor caso, em milisegundos, para os limiares 1, 2 e 3, do algoritmo de descarte de características da PMF dos objetos.

Limiar da PMF	Tempo Médio de Classificação no Melhor Caso (ms)
1	281
2	270
3	265

6 Características Temporais

Para estimar as características temporais dos algoritmos desenvolvidos, foram testadas três bibliotecas para medição de tempo utilizando as linguagens de programação C e C++, e escolheu-se a `timeb.h` (IEEE, 2013a) que é uma biblioteca para medição de tempo definida pelo padrão POSIX (*Portable Operating System Interface*) e implementada pelos compiladores de C compatíveis com esse padrão. Em seguida, são apresentados os tempos médios tanto de localização das amostras de teste no melhor caso, quanto de construção dos mapas.

6.1 Métricas para Medição de Tempo

Antes de efetuar as medições de tempo dos algoritmos desenvolvidos, foram feitos testes com as bibliotecas implementadas em C e C++, `time.h` (IEEE, 2013b), `timeb.h` (IEEE, 2013a) e `timer.hpp` (DAWES, 2013), com o objetivo de comprovar a resolução da medição de tempo fornecida por cada uma destas. Os testes efetuados consistiram em medir o tempo de execução de um trecho de código genérico pelas três bibliotecas e comparar os resultados, como apresentado na Figura 6.1.

Apesar do sistema operacional utilizado ser multitarefa, Ubuntu 12.04 (UBUNTU, 2013), e por isso o tempo medido não ser exatamente o tempo gasto pelo trecho de código definido, foi possível descartar as bibliotecas `time.h` e `timer.hpp`, porque a primeira apresentava resolução de apenas 1 segundo e a segunda mostrou-se instável, apesar da resolução nominal de 1 milissegundo. Logo, optou-se pela utilização da biblioteca `timeb.h`, por ter apresentado resolução real de 1 milissegundo e comportamento estável.

```

1  #include <time.h>
2  #include <sys/timeb.h>
3  #include <boost/timer.hpp>
4  #include <iostream>
5
6  using namespace std;
7
8  typedef struct timeb time_ref;
9
10 int main()
11 {
12     time_ref aux1, aux2;
13     unsigned long long int i;
14     unsigned long long int n;
15     time_t init, end;
16     boost::timer aux_timer;
17     double aux_t;
18
19
20     cout << "Forneca n:";
21     cin >> n;
22     cout << n << endl;
23
24     ftime(&aux1);
25     for(i=0; i < n; i++);
26     ftime(&aux2);
27
28     cout << "tempo da biblioteca timeb: " << 1000*(aux2.time-aux1.time) + (aux2.millitm -
        aux1.millitm) << " ms" << endl;
29
30     init = time(NULL);
31     for(i=0; i < n; i++);
32     end = time(NULL);
33
34     cout << "tempo da biblioteca time: " << 1000*(end-init) << " ms" << endl;
35
36     aux_timer.restart();
37     for(i=0; i < n; i++);
38     aux_t = aux_timer.elapsed();
39
40     cout << "tempo da biblioteca boost_1_40: " << 1000*aux_t << " ms" << endl;
41
42     return 0;
43 }

```

Figura 47 – Código de teste das bibliotecas de medição de tempo.

6.2 Análise de Tempo de Resposta

Uma vez definida a biblioteca a ser utilizada para medição de tempo, foram definidas as seguintes características temporais a serem medidas nos algoritmos desenvolvidos: o tempo de localização ou reconhecimento das amostras de teste; e o tempo de construção dos mapas.

Para todas as execuções dos algoritmos foi utilizado o sistema operacional Ubuntu 12.04 (UBUNTU, 2013), e as máquinas utilizadas foram três notebooks com configurações diferentes, que são descritas em detalhes nos Anexos A, A e A.

Vale ressaltar que todos os testes apresentados neste capítulo, e o cálculo de seus respectivos tempos de execução, foram executados no Notebook 3, Anexo A. No entanto, não se pode atribuir os tempos de localização alcançados nos teste à esta máquina que apresenta alto poder de processamento¹, por causa de algumas características da implementação dos algoritmos, descritas abaixo.

A biblioteca do método SURF que foi utilizada (ETH, 2013), implementa a extração das características em memória e pode manter os descritores em memória ou retorná-los para arquivos. Contudo, para evitar problemas de capacidade de armazenamento ao aplicar o extrator SURF em ambientes externos, já previamente apontados por Cummins e Newman (2008) e Cummins e Newman (2010), optou-se por armazenar os descritores SURF em arquivo.

Sendo assim, a maior limitação em relação às características temporais dos algoritmos implementados está na velocidade de acesso do disco rígido utilizado. Justamente por isso, utilizou-se a máquina notebook 3 para executar os testes.

Para comprovar esta hipótese, analisou-se a utilização dos recursos do sistema durante a execução dos teste através da aplicação monitor do sistema, do Ubuntu 12.04, e nenhum gráfico de utilização do processador alcançou 150%. Ou seja, durante todo o tempo de execução dos algoritmos, tanto de mapeamento quanto de localização, não utilizou-se mais do que o equivalente a um processador e meio, como mostra a Figura 6.2.

¹ Considera-se que o notebook 3 apresenta alto poder de processamento por apresentar um processador com 4 núcleos e 6 GB de *Random Access Memory* (RAM).

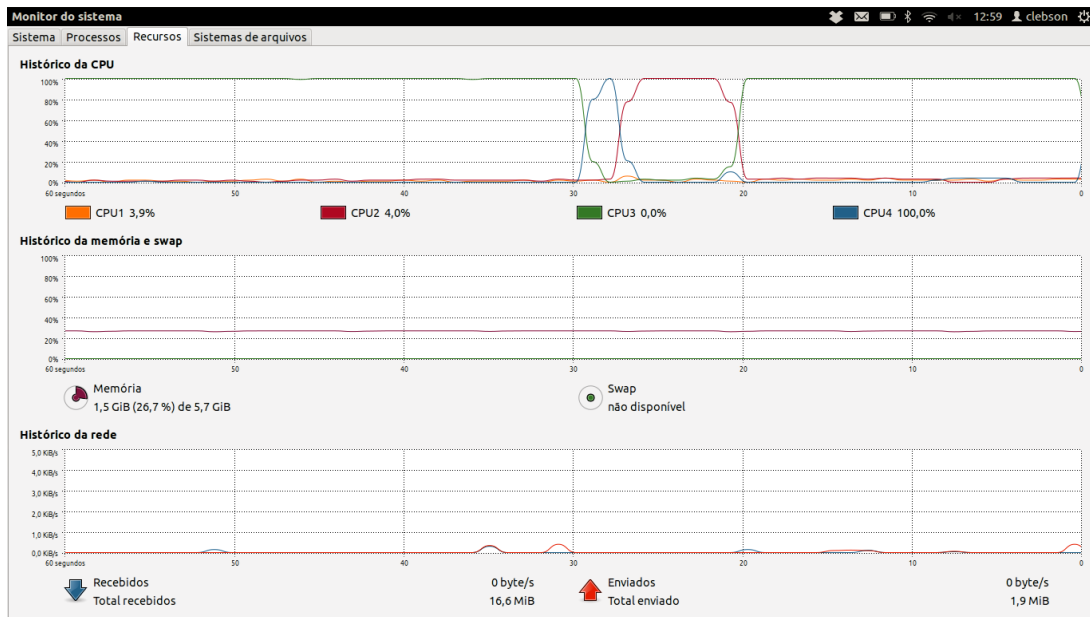


Figura 48 – Gráfico de utilização dos recursos do sistema durante a execução dos algoritmos de mapeamento de localização. Para todos os teste executados, os gráficos gerados pelo monitor do sistema apresentaram as mesma características, tais quais as apresentadas nesta figura.

Então, pode-se concluir que os algoritmos implementados podem ser aplicados a ambientes externos com poucas restrições em relação à capacidade da memória RAM, pois a limitação de armazenamento será definida pelo disco rígido utilizado, que normalmente apresenta capacidade de armazenamento muito superior à da memória RAM.

6.3 Tempos de Resposta dos Experimentos

Para finalizar a discussão sobre as características temporais dos algoritmos, são apresentados na Tabela 17, os tempos gastos nas construções dos mapas e nas localizações das amostras de teste, no melhor caso, dos experimentos apresentados anteriormente.

Tabela 17 – Tempo de médio de localização/classificação das amostras de teste no melhor caso, e de construção do mapa para os principais experimentos apresentados no Capítulo 5.

Experimento	Limiar de Descarte	Nº de Lugares/Objetos	Construção do Mapa (min)	Localização no Melhor Caso (s)
Figura 28	1	5	1,6912	1,448
Figura 32	1	27	111,9480	26,082
Amostras da Figura 32	10	27	109,1407	6,129
Figura 36	1	28	111,3972	12,283
Amostras da Figura 36	10	28	118,3153	3,361
Figura 45	1	4	0,0609	0,281
Amostras da Figura 45	3	4	0,0579	0,265

Pela Tabela 17, percebe-se que o tempo de construção do mapa para os experimentos de localização em ambientes externos é alto para uma aplicação *online*. Porém, a proposta deste trabalho foi desenvolver uma método que contruísse o mapa de forma *offline*, e por isso considera-se que para este propósito o tempo de construção do mapa é adequado. Todavia, não descarta-se a possibilidade da implementação de melhorias nesse algoritmo, pois não foi utilizada nenhuma forma de paralelismo, através de *threads* por exemplo, na construção dos lugares, que podem ser facilmente paralelizados pois não dependem entre si.

Com respeito ao algoritmo de localização, também podem ser efetuadas melhorias para reduzir o tempo de localização, tais como a paralelização da extração de características das imagens, e a elaboração de uma abordagem mista com relação ao armazenamento dos descritores SURF. Assim, pode-se carregar parte destes em memória de forma dinâmica, ao analisar a demanda do método de localização, que nem sempre utilizará todos os lugares no processo de decisão.

Além disso, pode-se melhorar o tempo de localização ao utilizar uma estratégia de navegação construtiva ou topológica do mapa, que considere a ligação entre os nós e/ou o conhecimento *a priori*.

7 Conclusão

No contexto da robótica móvel em ambientes externos e dinâmicos, nota-se a existência de alguns problemas ainda não resolvidos por completo (SPERO, 2004; KYRKI; KRAGIC, 2011), tais como: mapeamento, planejamento de trajetória, localização, navegação e identificação de objetos que apresentem certas características.

Como muitos desses problemas apresentam chances de serem resolvidos por Visão Computacional, e o problema de localização é considerado como um dos principais problemas em robótica, o objetivo deste trabalho foi desenvolver um método de localização visual autônoma *online*, sem considerar informações *a priori*, que possibilitasse uma forma de navegação robótica visual autônoma *online* em ambientes externos e dinâmicos, baseada no método de extração e descrição de características visuais SURF (BAY *et al.*, 2008), aplicado a imagens panorâmicas obtidas por uma única câmera perspectiva.

Como a abordagem escolhida não utilizou informações *a priori* com relação à posição da câmera, não foi possível alcançar baixos tempos de localização das amostras no pior caso. Todavia, se forem incorporadas informações de localização ao longo da navegação e a topologia do mapa, será possível reduzir significativamente o tempo de localização das amostras.

Assim, desenvolveu-se o modelo para caracterização visual, apresentado no Capítulo 3, que se baseou na estimação das PMFs de características visuais SURF para os lugares do mapa. Utilizando esse modelo para localizar amostras de lugares do mapa, definiu-se um método de localização, apresentado na Equação 3.9, que calcula as probabilidades de uma amostra pertencer a lugares candidatos do mapa, e as compara a valores de referência, definidos por curvas ROC dos lugares candidatos.

Em seguida, foram realizados testes para avaliar o poder de caracterização do modelo desenvolvido e a qualidade do método de localização proposto. Para isso, aplicou-se este modelo ao problema da localização de amostras visuais em um ambiente externo dinâmico, composto por 28 lugares; e ao problema de reconhecimento de objetos em um conjunto formado por 4 objetos similares.

Frente aos resultados apresentados nos Capítulos 5 e 6, comprova-se que este trabalho alcançou seus objetivos, pois amostras visuais de um ambiente externo e dinâmico foram localizadas corretamente, alcançando desempenho global pelo menos aceitável ($AUC > 0,7$), gastando 3,361 segundos no melhor caso; e também foi possível reconhecer objetos do conjunto utilizado também com desempenho global pelo menos aceitável, gastando 265 milisegundos no melhor caso.

Por fim, observando o potencial de aplicação do modelo proposto são sugeridos os seguintes trabalhos:

- desenvolvimento de um modelo de normalização global das PMFs dos lugares, pois observou-se que um dos principais problemas do modelo proposto foi a diferença do número de características nas PMFs dos lugares do mapa;
- realização das estimativas das PMFs pelos métodos de inferência por Janelas de Parzen e K-NN;
- desenvolvimento de um algoritmo que selecione um pequeno conjunto de características dentre as extraídas dos lugares, para que apenas estas sejam utilizadas para caracterizá-los, diminuindo assim o tempo de localização das amostras;
- escolha/desenvolvimento de filtros para serem aplicados às imagens na etapa de pré-processamento, com o objetivo de reduzir a quantidade de características extraídas do ambiente e consequentemente reduzir tanto o tempo de construção do mapa, quanto o tempo de localização das amostras;
- alteração da implementação do modelo, com o objetivo de paralelizar a criação das PMFs dos lugares, e consequentemente diminuir o tempo de construção do mapa;
- desenvolvimento e implementação de uma política de armazenamento das características, descritores SURF, que defina dinamicamente se estas devem ser armazenadas em disco ou em memória;
- aplicação do modelo desenvolvido a um robô móvel, com o objetivo de validar os resultados obtidos neste trabalho;
- aplicar o modelo desenvolvido a um conjunto de objetos maior, com o objetivo de validar os resultados obtidos neste trabalho;
- desenvolvimento de um modelo de navegação para robôs móveis, que utilize câmeras, através do modelo desenvolvido, e sensores baseados em GPS para permitir a estes robôs navegarem de forma robusta em ambientes externos dinâmicos;
- desenvolver uma aplicação para dispositivos pessoais móveis, tais como *smartphones* e *tablets*, que permita a criação de mapas personalizados para as demandas de localização de cada pessoa. Um exemplo seria a criação de mapas em *shopping centers* para ajudar as pessoas a se localizarem.

Referências

- ABB. *Robôs*. 2012. Acessado em 02 de fevereiro às 10:29. Disponível em: <http://www.abb.com/product/us/9AAC100735.aspx>. Citado na página 20.
- ADOLF, F. *Cleaning & Housekeeping Robots*. 2012. Acessado em 02 de fevereiro às 12:54. Disponível em: <http://www.service-robots.org/applications/cleaning.htm>. Citado na página 20.
- ANDREASSON, H.; DUCKETT, T. Topological localization for mobile robots using omni-directional vision and local features. In: *In Proceedings of the 5th IFAC Symposium on Intelligent Autonomous Vehicles*. [S.l.: s.n.], 2004. Citado na página 26.
- APPLE. *O que é o iCloud?* 2012. Acessado em 31 de janeiro às 15:48. Disponível em: <http://www.apple.com/br/icloud/what-is.html>. Citado na página 20.
- ARMBRUST, M.; FOX, A.; GRIFFITH, R.; JOSEPH, A. D.; KATZ, R.; KONWINSKI, A.; LEE, G.; PATTERSON, D.; RABKIN, A.; STOICA, I.; ZAHARIA, M. *Above the Clouds: A Berkeley View of Cloud Computing*. 2009. Acessado em 31 de janeiro às 15:31. Disponível em: <http://radlab.cs.berkeley.edu/>. Citado na página 20.
- BAY, H.; ESS, A.; TUYTELAARS, T.; GOOL, L. V. SURF: Speeded up robust features. *Computer Vision and Image Understanding (CVIU)*, v. 110, n. 3, p. 346–359, 2008. Citado 19 vezes nas páginas 9, 10, 26, 28, 30, 31, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44 e 127.
- BAY, H.; TUYTELAARS, T.; GOOL, L. V. SURF: Speeded up robust features. In: *9th European Conference on Computer Vision*. [S.l.: s.n.], 2006. Citado na página 26.
- BOADA, B. L.; BLANCO, D.; MORENO, L. Symbolic place recognition in voronoi-based maps by using hidden markov models. *Journal of Intelligent and Robotic Systems*, Springer Netherlands, v. 39, p. 173–197, 2004. Citado na página 25.
- BROWN, M.; LOWE, D. Invariant features from interest point groups. In: *BMVC 2002: 13th British Machine Vision Conference*. [s.n.], 2002. p. 253–262. Disponível em: <http://opus.bath.ac.uk/26128/>. Citado na página 36.
- CHEN, Y. *Design and Evaluation of an Autonomous Vehicle Based on GPS Navigation*. Dissertação (Mestrado) — Uppsala University, Department of Information Technology, 2013. Citado 2 vezes nas páginas 22 e 23.
- CUMMINS, M.; NEWMAN, P. Fab-map: Probabilistic localization and mapping in the space of appearance. *The International Journal of Robotics Research*, v. 27, n. 6, p. 647–665, 2008. Disponível em: <http://ijr.sagepub.com/content/27/6/647.abstract>. Citado 2 vezes nas páginas 27 e 124.
- CUMMINS, M.; NEWMAN, P. Appearance-only slam at large scale with fab-map 2.0. *The International Journal of Robotics Research*, 2010. Disponível em: <http://ijr.sagepub.com/content/early/2010/11/11/0278364910385483.abstract>. Citado 3 vezes nas páginas 27, 44 e 124.

- CôCO, K. F. *Modelos Fractais para a Função de Vizinhança na Análise Topográfica de Componentes Independentes*. Tese (Doutorado) — PPGE/ UFES, 2007. Orientador(es): Evandro Ottoni Teatini Salles, Mario Sarcinelli Filho. Citado na página 30.
- DAVIS, J.; GOADRIC, M. The relationship between precision-recall and roc curves. In: *Proceedings of the 23rd international conference on Machine learning*. [S.l.: s.n.], 2006. p. 233–240. Citado 2 vezes nas páginas 42 e 43.
- DAWES, B. *Timer Home*. 2013. Acessado em 03 de janeiro às 12:00. Disponível em: <<http://www.bost.org/doc/libs/1520/boost/timer.hpp>>. Citado na página 122.
- DUDEK, G.; JUGESSUR, D. Robust place recognition using local appearance based methods. In: *IEEE Intl. Conf. on Robotics and Automation (ICRA)*. [S.l.]: IEEE Press, 2000. p. 1030–1035. Citado na página 25.
- ETH. *SURF Download*. 2013. Acessado em 03 de janeiro às 14:28. Disponível em: <<http://www.vision.ee.ethz.ch/surf/downloadac.html>>. Citado na página 124.
- FAWCETT, T. *An introduction to ROC analysis*. [S.l.]: Elsevier, 2006. 861-874 p. Citado 5 vezes nas páginas 11, 43, 60, 61 e 62.
- FERGUS, R.; PERONA, P.; ZISSERMAN, A. Object class recognition by unsupervised scale-invariant learning. *2012 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, Los Alamitos, CA, USA, v. 2, p. 264, 2003. ISSN 1063-6919. Citado na página 43.
- GAUGLITZ, S.; HÖLLERER, T.; TURK, M. Evaluation of interest point detectors and feature descriptors for visual tracking. *International Journal of Computer Vision*, Springer Netherlands, v. 94, p. 335–360, 2011. ISSN 0920-5691. 10.1007/s11263-011-0431-5. Disponível em: <<http://dx.doi.org/10.1007/s11263-011-0431-5>>. Citado na página 31.
- GRISSE, R.; ALLEY, M. M.; HEATWOLE, C. *Precision Farming Tools: Global Positioning System (GPS)*. [S.l.]: Virginia Cooperative Extension, 2009. Publication 442-503. Citado na página 23.
- HAZAS, M.; SCOTT, J.; KRUMM, J. Location-aware computing comes of age. *Computer*, v. 37, n. 2, p. 95 – 97, feb 2004. ISSN 0018-9162. Citado 2 vezes nas páginas 9 e 22.
- IEEE, O. G. *sys/timeb.h - additional definitions for date and time*. 2013. Acessado em 03 de janeiro às 11:48. Disponível em: <<http://pubs.opengroup.org/onlinepubs/009695399/basedefs/sys/timeb.h.html>>. Citado na página 122.
- IEEE, O. G. *time.h - time types*. 2013. Acessado em 03 de janeiro às 11:49. Disponível em: <<http://pubs.opengroup.org/onlinepubs/009696699/basedefs/time.h.html>>. Citado na página 122.
- JENSEN, K.; LARSEN, M.; SIMONSEN, T.; JORGENSEN, R. N. Evaluating the performance of a low-cost gps in precision agriculture applications. *First RHEA International Conference on Robotics and associated High-technologies and Equipment for Agriculture*, 2012. Citado na página 22.
- KOENDERINK, J. The structure of images. *Biological Cybernetics*, Springer Berlin / Heidelberg, v. 50, p. 363–370, 1984. ISSN 0340-1200. 10.1007/BF00336961. Disponível em: <<http://dx.doi.org/10.1007/BF00336961>>. Citado na página 33.

- KYRKI, V.; KRAGIC, D. Computer and robot vision. *IEEE Robotics Automation Magazine*, v. 18, p. 121–122, 2011. Citado 2 vezes nas páginas 20 e 127.
- LINDEBERG, T. Scale-space for discrete signals. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, v. 12, n. 3, p. 234–254, mar 1990. ISSN 0162-8828. Citado 2 vezes nas páginas 33 e 34.
- LOWE, D. Object recognition from local scale-invariant features. In: *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*. [S.l.: s.n.], 1999. v. 2, p. 1150–1157 vol.2. Citado 2 vezes nas páginas 26 e 36.
- LOWE, D.; LITTLE, J. Local and global localization for mobile robots using visual landmarks. In: *International Conference on Intelligent Robots and Systems (IROS 01)*. [S.l.: s.n.], 2001. Citado na página 26.
- LOWE, D. G. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60, Kluwer Academic Publishers, p. 91–110, 2004. Citado 2 vezes nas páginas 26 e 41.
- MELL, P.; GRANCE, T. *The NIST Definition of Cloud Computing (Draft) - Recommendations of the National Institute of Standards and Technology*. 2011. Special Publication 800-145 (Draft). Acessado em 31 de janeiro às 15:45. Citado na página 20.
- MIKOLAJCZYK, K.; SCHMID, C. Scale & affine invariant interest point detectors. *Int. J. Comput. Vision*, Kluwer Academic Publishers, Hingham, MA, USA, v. 60, n. 1, p. 63–86, out. 2004. ISSN 0920-5691. Disponível em: <<http://dx.doi.org/10.1023/B:VISI.0000027790.02288.f2>>. Citado na página 41.
- MILLER, S. L.; CHILDERS, D. G. *Probability and Random Processes: With Applications To Signal Processing And Communications*. [S.l.]: Elsevier Academic Press, 2004. Citado na página 48.
- MITSUHASHI, M.; KURODA, Y. Mobile robot localization using place recognition in outdoor environments with similar scenes. In: *Advanced Intelligent Mechatronics (AIM), 2011 IEEE/ASME International Conference on*. [S.l.: s.n.], 2011. p. 930–935. ISSN 2159-6247. Citado 2 vezes nas páginas 27 e 44.
- MITSUHASHI, M.; SAKAI, A.; KURODA, Y. Mobile robot localization using appearance based place recognition. In: *Robotics and Biomimetics (ROBIO), 2010 IEEE International Conference on*. [S.l.: s.n.], 2010. p. 537–542. Citado na página 27.
- NASA, J. P. L. *ATHLETE - All-Terrain Hex-Limbed Extra-Terrestrial Explorer*. 2012. Acessado em 22 de julho às 14:55. Disponível em: <<http://athlete.jpl.nasa.gov/>>. Citado 2 vezes nas páginas 9 e 24.
- OCEANEERING. *Remotely Operated Vehicles (ROVs)*. 2012. Acessado em 02 de fevereiro às 12:41. Disponível em: <<http://www.oceaneering.com/rovs/>>. Citado na página 20.
- OLIVEIRA, C. J. M. de; VASSALLO, R. F.; SALLES, E. O. T. Visual outdoor localization using a probabilistic model. In: *Anais do X Simpósio Brasileiro de Automação Inteligente (XSBAI)*. [S.l.: s.n.], 2011. v. 10, p. 1250–1255. Citado 2 vezes nas páginas 66 e 68.

- ROBOTSHOP. *Personal and Domestic Robots*. 2012. Acessado em 02 de fevereiro às 10:33. Disponível em: <<http://www.robotshop.com/personal-domestic-robots.html>>. Citado na página 20.
- SANCHEZ-GONZALEZ, L.; GARCIA, F.; RUIZ, F.; MENDLING, J. A study of the effectiveness of two threshold definition techniques. In: *Evaluation Assessment in Software Engineering (EASE 2012), 16th International Conference on*. [S.l.: s.n.], 2012. p. 197–205. Citado 2 vezes nas páginas 57 e 62.
- SCHMID, C.; MOHR, R.; BAUCKHAGE, C. Evaluation of interest point detectors. *Int. J. Comput. Vision*, Kluwer Academic Publishers, Hingham, MA, USA, v. 37, n. 2, p. 151–172, jun. 2000. ISSN 0920-5691. Disponível em: <<http://dx.doi.org/10.1023/A:1008199403446>>. Citado na página 41.
- SPERO, D. J. *A Review of Outdoor Robotics Research - MECSE-17-2004*. [S.l.], 2004. Disponível em: <<http://www.dorianspero.com/Spero04a.pdf?attredirects=0>>. Citado 4 vezes nas páginas 20, 24, 31 e 127.
- TAKASU, T.; YASUDA, A. Development of the low-cost rtk-gps receiver with an open source program package rtklib. In: *International Symposium on GPS/GNSS, International Convention Center Jeju, Korea*. [S.l.: s.n.], 2009. Citado na página 23.
- THRUN, S.; BURGARD, W.; FOX, D. *Probabilistic Robotics*. [S.l.]: Massachusetts Institute of Technology, 2005. Citado 4 vezes nas páginas 9, 21, 25 e 47.
- TORRALBA, A.; OLIVA, A. Statistics of natural image categories. *Network: computation in neural systems*, Informa UK Ltd UK, v. 14, n. 3, p. 391–412, 2003. Citado na página 30.
- UBUNTU. *Ubuntu 12.04.1 LTS (Precise Pangolin)*. 2013. Acessado em 03 de janeiro às 13:32. Disponível em: <<http://releases.ubuntu.com/precise/>>. Citado 2 vezes nas páginas 122 e 124.
- ULRICH, I.; NOURBAKHSH, I. Appearance-based place recognition for topological localization. In: *Robotics and Automation, 2000. Proceedings. ICRA '00. IEEE International Conference on Robotics and Automation*. [S.l.: s.n.], 2000. v. 2, p. 1023–1029 vol.2. Citado na página 25.
- VALGREN, C.; LILIENTHAL, A. Incremental spectral clustering and seasons: Appearance-based localization in outdoor environments. In: *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*. [S.l.: s.n.], 2008. Citado na página 26.
- VIOLA, P.; JONES, M. Rapid object detection using a boosted cascade of simple features. In: *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*. [S.l.: s.n.], 2001. v. 1, p. I–511 – I–518 vol.1. ISSN 1063-6919. Citado na página 32.
- ZHANG, S.; XIAO, W.; XIE, L. *A Novel Feature Extraction Algorithm for Outdoor Mobile Robot Localization*. [S.l.]: ISBN: 978-953-7619-01-5, InTech, 2008. Citado na página 24.

Anexos

ANEXO A – Equipamentos Utilizados

Câmera

- Fabricante: MOBOTIX;
- Modelo: M24M.



Figura 49 – Câmera.

Notebook 1

- Fabricante: HP;
- Modelo: Pavilion.
 - Processador: Core Duo 1,86 GHz;
 - Memória: 1 GB;
 - Velocidade do HD: 5600 RPM.



Figura 50 – Notebook 1.

Notebook 2

- Fabricante: Dell;

- Modelo: Vostro 3550.
 - Processador: Core i5 2,53 GHz;
 - Memória: 4 GB;
 - Velocidade do HD: 5600 RPM.



Figura 51 – Notebook 2.

Notebook 3

- Fabricante: Dell;
- Modelo: Vostro 3450.
 - Processador: Core i7 2,8 GHz;
 - Memória: 6 GB;
 - Velocidade do HD: 7200 RPM.

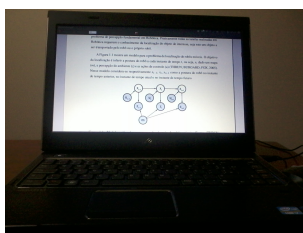


Figura 52 – Notebook 3.